DOCUMENT RESUME

ED 093 716 SE 018 093

AUTHOR Wiles, Clyde A.
TITLE Comparison of Three Instructional Sequences for the Addition and Subtraction Algorithms. Technical Report 273.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO TR-273
PUB DATE Aug 73
CONTRACT NE-C-00-3-0065; OEC-5-10-154
NOTE 277p.

EDRS PRICE MF-$0.75 HC-$13.80 PLUS POSTAGE
DESCRIPTORS *Algorithms; Doctoral Theses; *Elementary School Mathematics; Grade 2; *Instruction; Mathematics Education; *Research

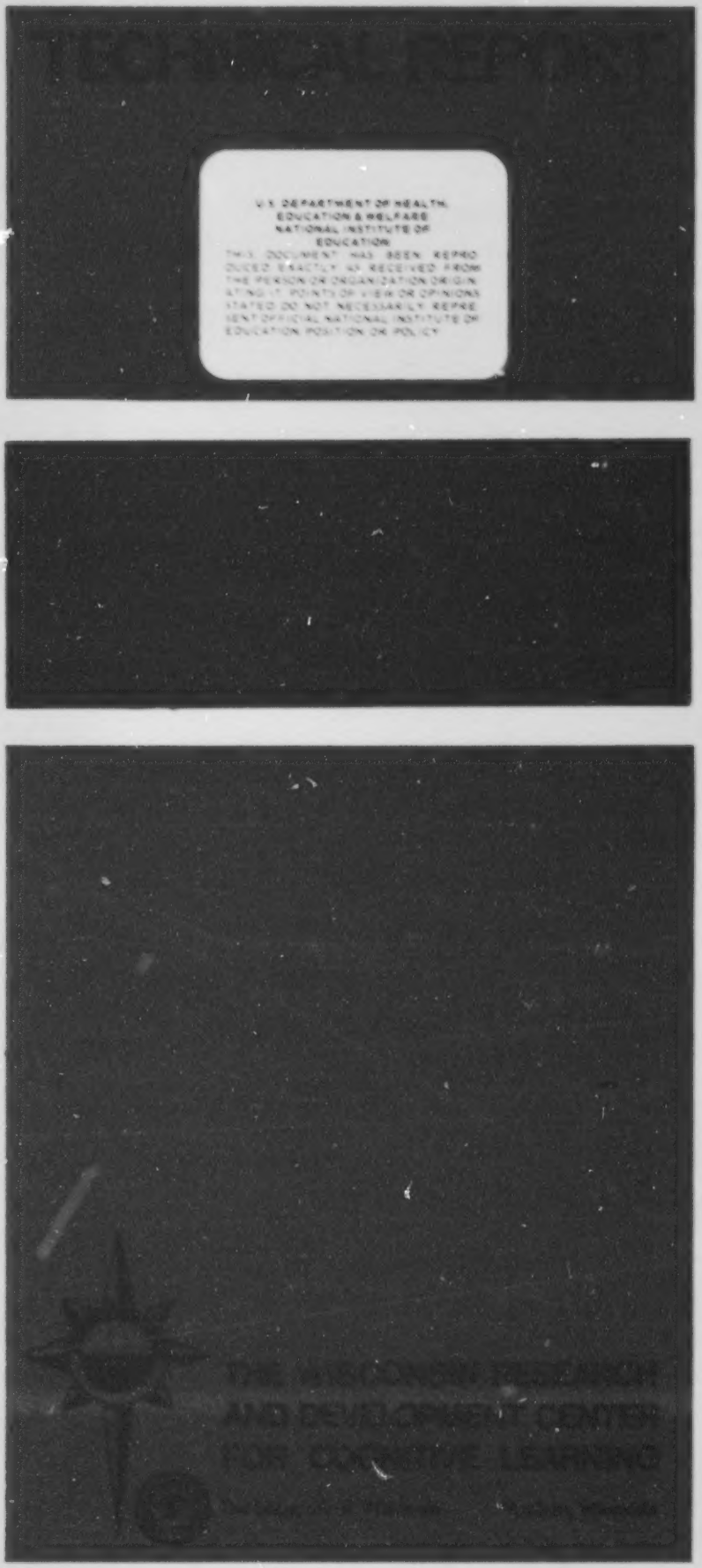
ABSTRACT

The study's purpose was to investigate the differential effects on the achievement of second-grade students that could be attributed to three instructional sequences for the learning of the addition and subtraction algorithms. One sequence presented the addition algorithm first (AS), the second presented the subtraction algorithm first (SA), and the third was an integrated presentation of the two algorithms (I). Students were randomly assigned from IQ blocks to the three treatment groups. The performance of 49 subjects was measured prior to instruction, during instruction, immediately following six weeks of instruction, and two weeks after instruction; posttest and retention test included both timed and power test variables. Data were analyzed using multivaraite analysis of covariance. Findings showed: (1) the instructional sequence SA produced generally lower performance than either of the other two treatments; the tendency of all groups was to learn addition first, (2) the amount of learning and retention measured by the timed variables was significantly greater for group AS than for either of the other two groups, and (3) on the retention test, except for addition facts and addition without regrouping, the overall performance of group I for the timed variables was significantly greater than that of group SA. (DT)

ED 093716

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SE



SE 018 093

U.S. Office of Education Center No. C-03 Contract OE 5-10-154

ED 093716

Technical Report No. 273

COMPARISON OF THREE INSTRUCTIONAL SEQUENCES FOR THE ADDITION AND SUBTRACTION ALGORITHMS

Report from the Project on Conditions of School Learning and Instructional Strategies

by Clyde A. Wiles

John G. Harvey and Thomas A. Romberg
Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

August 1973

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred

Center Contract No. NE-C-00-3-0065

STATEMENT OF FOCUS

Individually Guided Education (IGE) is a new comprehensive system of elementary education. The following components of the IGE system are in varying stages of development and implementation: a new organization for instruction and related administrative arrangements; a model of instructional programing for the individual student; and curriculum components in prereading, reading, mathematics, motivation, and environmental education. The development of other curriculum components, of a system for managing instruction by computer, and of instructional strategies is needed to complete the system. Continuing programmatic research is required to provide a sound knowledge base for the components under development and for improved second generation components. Finally, systematic implementation is essential so that the products will function properly in the IGE schools.

The Center plans and carries out the research, development, and implementation components of its IGE program in this sequence: (1) identify the needs and delimit the component problem area; (2) assess the possible constraints--financial resources and availability of staff; (3) formulate general plans and specific procedures for solving the problems; (4) secure and allocate human and material resources to carry out the plans; (5) provide for effective communication among personnel and efficient management of activities and resources; and (6) evaluate the effectiveness of each activity and its contribution to the total program and correct any difficulties through feedback mechanisms and appropriate management techniques.

A self-renewing system of elementary education is projected in each participating elementary school, i.e., one which is less dependent on external sources for direction and is more responsive to the needs of the children attending each particular school. In the IGE schools, Center-developed and other curriculum products compatible with the Center's instructional programing model will lead to higher morale and job satisfaction among educational personnel. Each developmental product makes its unique contribution to IGE as it is implemented in the schools. The various research components add to the knowledge of Center practitioners, developers, and theorists.

ACKNOWLEDGEMENTS

It is not possible to adequately thank all who have contributed to the completion of this thesis, however, the author wishes to publicly acknowledge the contributions of the following people:

John G. Harvey, thesis advisor, and the other members of the reading committee, Thomas A. Romberg and J. Fred Weaver, all of whom contributed liberally of their time and efforts;

The staff and children of Stephen Bull Elementary School in Racine, Wisconsin, and, in particular, the teachers; Patricia Hansen, Marlene Schiestle, and Janis Zinns who made it possible to conduct this study;

The staff of the Wisconsin Research and Development Center for Cognitive Learning who aided in the preparation of materials, data analysis, and the preparation of the manuscript;

My wife, Barbara, and our children who have contributed encouragement, understanding, and patience without which this thesis could not have been undertaken or completed.

TABLE OF CONTENTS

		page
ACKNOWLEDGEMENTS		iv
LIST OF TABLES		xi
LIST OF FIGURES		xv
ABSTRACT		xvii
CHAPTER		
I	INTRODUCTION	1
	The Nature of an Algorithm	2
	The Role of Instructional Sequence	5
II	RESEARCH RELATING TO THE ALGORITHMS	9
	The Selection of an Algorithm	9
	The Method of Instruction	11
	The Dependent Variables Examined	13
III	RESEARCH WITH INSTRUCTIONAL SEQUENCES	17
	General Classification of Studies	17
	The Use of Integrated Sequences	19
	An Earlier Study	25
IV	SYSTEMS FOR CONSTRUCTING INSTRUCTIONAL SEQUENCES	31
	Definition of Sequence	31
	Systems Approaches to Curriculum Construction	32
	Systems Approaches to Instructional Sequence	34
	Summary	37

CHAPTER		page
V	THE PROBLEM AND THE SEQUENCE SELECTION MODEL USED	39
	The Analysis Phase	39
	Instructional Activities	41
	Definitions	43
	The Problem	45
VI	DESIGN AND PROCEDURES	49
	Introduction	49
	Subjects	50
	Treatments	52
	The Construction of the Activities	52
	Pacing of Instructional Activities	54
	Group I Management	55
	Management of Groups AS and SA	55
	Observations	56
	The Observation 01	56
	Achievement Monitoring Observations	57
	The Observation 02	59
	The Observation 03	59
	Administration and Scoring	60
	Analysis	60
VII	CONDUCT OF THE STUDY	61
	Subjects	61
	The Administration of the Experiment	64
	The Administration of Tests	69
	Exigencies of the Study	72
VIII	DATA DESCRIPTION AND ANALYSIS	75
	The Hypotheses	75
	The Specification of Variables	77
	Observations	77
	Status Variables	82

CHAPTER

	page
Periodic Observation Achievement Variables	82
Achievement Variables of Observations 01, 02, and 03	84
Power Test Variables of Observations 02 and 03	86
The Data Analysis Model Employed	90
Significance Level	92
The Status Variables and Observation 01 Data	93
The Periodic Achievement Measures	97
Summary of the Periodic Observation Data	116
Sequence of Data Analyses for 02 and 03 Data	118
The Selection of Covariates	120
The Covariates for the Achievement Variables	125
The Covariates for the Power Test Variables	125
The Posttest Data from Observation 02	128
The 02 Achievement Variables	128
Summary of 02 Achievement Variable Comparisons	136
The 02 Total Power Test Variables	137
The 02 Difference Power Test Variables	140
Summary of 02 Difference Power Test Variable Comparisons	149
Retention Test Data from Observation 03	151
The 03 Achievement Variables	152
Summary of 03 Achievement Variables Comparisons	162
The 03 Total Power Test Variables	163
The 03 Difference Power Test Variables	167
Summary of 03 Difference Power Test Variable Comparisons	173

CHAPTER

03 - 02 Data 173

The 03 - 02 Achievement Variables . . . 177
Summary: 03 - 02 Achievement Test Variables. 179
The 03 - 02 Total Power Test Variables. 181
Summary: 03 - 02 Total Power Test Variables. 185
The 03 - 02 Difference Power Test Variables. 185
Summary of 03 - 02 Difference Power Test Variable Comparisons. 192

Summary of the Data. 192

IX CONCLUSIONS 201

The Findings of the Study. 201

Findings Related to Achievement Monitoring Data 201
Findings Related to Observations 02 and 03 202

Limitations. 205

The Quality of the Instructional Materials. 205
Limitations of the Tests. 207
Limitations of the Analysis 208
Confounding Factors 209
Limitations to Generalizability 212
The Type of Materials Used. 213
The Population. 213

Discussion of Results and Implication for Sequencing. 214

Summary 220

Implications for Research. 223

The Identification of Coordinate Tasks. 223
The Specification of Tasks and Variables 225
The Learning of Algorithms. 226

Concluding Remarks 228

page

REFERENCES 231

APPENDIX A BEHAVIORS ASSOCIATED WITH THE ADDITION AND SUBTRACTION ALGORITHMS FOR TWO-DIGIT WHOLE NUMBERS 241

APPENDIX D DATA ON SUBJECTS IN THE EXPERIMENT. 249

LIST OF TABLES

TABLE		page
1	The Age Distribution of Subjects	62
2	The IQs of Subjects by Sex, Block, and Group Assignment	65
3	The Reassignment of Teachers on the 15th Instructional Day.	68
4	An Overview of the Conduct of the Study. . . .	70
5	Missing Data for Each Group at Each Observation.	73
6	The Observations and Variables	78
7	Observed Group Means for 01 Variables and Status Variables	94
8	Estimated Group Means and Standard Errors of 01 Variables and Status Variables	94
9	MANOVA and ANOVAs for 01 Variables	96
10	Observed Group Means and Standard Deviations of Variables at the 12 Periodic Observations .	90
11	Total Group Means for the 12 Periodic Observations	105
12	Estimated Group Means and Standard Errors for Four Time Periods.	109
13	MANOVAs and ANOVAs of Group Means for T Variables.	114
14	F Tests of No Association Between the Achievement Variables of 02 and 03 and the Potential Covariates	126
15	F Tests of No Association Between the Power Test Variables of 02 and 03 and the Potential Covariates	127

TABLE		page
16	Estimated Group Means and Standard Errors for 02 Variables	129
17	Adjusted Group Means for the 02 Achievement Variables.	132
18	Overall Groups MANCOVA and ANCOVAs for 02 Achievement Variables.	132
19	Discriminant Analysis: 02 Achievement Test Variables Adjusted Means	133
20	Pairs of Groups MANCOVAs and ANOVAs for 02 Achievement Variables.	138
21	Adjusted Group Means for 02 Total Power Test Variables.	139
22	Overall Groups MANCOVA and ANCOVA of 02 Total Power Test Variables	141
23	Pairs of Groups MANCOVA and ANCOVA for 02 Total Power Test Variables	142
24	Overall Groups MANOVA and ANOVAs for 02 Difference Power Test Variables.	145
25	Discriminant Analysis: 02 Difference Power Test Variables	146
26	Pairs of Groups MANOVA and ANOVA for 02 Difference Power Test Variables.	150
27	Estimated Group Means and Standard Errors for 03 Variables	153
28	Adjusted Group Means for the 03 Achievement Variables.	154
29	Overall Groups MANCOVA and ANCOVAs for 03 Achievement Variables.	156

TABLE		page
30	Discriminant Analysis: 03 Achievement Variables Adjusted Means	158
31	MANCOVAs and ANCOVAs for 03 Achievement Variables by Pairs of Groups	161
32	Adjusted Group Means for the 03 Total Power Test Variables	165
33	Overall Groups MANCOVA and ANCOVAs for 03 Total Power Test Variables	165
34	Pairs of Groups MANCOVAs and ANCOVAs for 03 Total Power Test Variables	166
35	Overall Groups MANOVA and ANOVAs for 03 Difference Power Test Variables.	169
36	Discriminant Analysis: 03 Variables Difference Power Test Variables.	170
37	Pairs of Groups MANCOVAs and ANCOVAs for Difference Power Test Variables.	174
38	Estimated Group Means for 03 - 02 Variables. .	176
39	Adjusted Group Means for 03 - 02 Achievement Variables.	178
40	Overall Groups MANCOVA and ANCOVAs for 03 - 02 Achievement Variables.	178
41	MANCOVAs and ANCOVAs for 03 - 02 Achievement Variables by Pairs of Groups	180
42	Adjusted Group Means for 03 - 02 Total Power Test Variables	183
43	Overall Groups MANCOVA and ANCOVAs for 03 - 02 Total Power Test Variables	183

TABLE		page
44	MANCOVAs and ANCOVAs for 03 - 02 Total Power Test Variables by Pairs of Groups.	184
45	Overall Groups MANOVA and ANOVAs for 03 - 02 Difference Power Test Variables.	187
46	Discriminant Analysis: 03 - 02 Difference Power Test Variables	188
47	MANOVAs and ANOVAs for 03 - 02 Difference Power Test Variables by Pairs of Groups	191
48	p Values of F Tests Involving 02 and 03 Variables.	194
49	Proportion of Facts Tests Showing Algorithmic Interference	219

LIST OF FIGURES

FIGURE		page
1	Experimental Design	50
2	Graphs of Group Means for P1ANC Through P12ANC.	100
3	Graphs of Group Means for P1SNB Through P12 SNB	101
4	Graphs of Group Means for P1C Through P12C.	102
5	Graphs of Group Means for P1B Through P12B.	103
6	Graphs of Total Performance of Groups P1 Through P12.	106
7	Graphs of Group Means for Four Time Periods of Instruction.	111
8A	Analysis of Posttest Data	121
8B	Hypothesis Testing for Each Branch of the Posttest Data	122
9	Graphs of Group Means for Some Achievement Variables	195
10	Graphs of Adjusted Group Means for +F and -F.	196

ABSTRACT

The purpose of this study was to investigate the differential effects on the achievement of second grade Ss that could be attributed to three instructional sequences for the learning of the addition and subtraction algorithms. The three sequences were designed so that one presented the addition algorithm first (AS), the second presented the subtraction algorithm first (SA), and the third was an integrated presentation of the two algorithms (I). The investigation sought answers to the following questions.

I. What differences attributable to instructional sequence can be found on measures of the following variables?
   1. The rate at which the algorithms are learned,
   2. the achievement levels for the terminal behaviors observed after instruction, and
   3. the retention of the terminal behaviors.

II. Do measures of proficiency with the addition and subtraction facts show significant changes due to the incidental drill inherent in learning the usual algorithms?

Information relating to these questions was gathered in the following way. The Ss were randomly assigned from IQ blocks to the three treatment groups. The performance of 49 subjects was measured

prior to instruction, 01, during instruction, P, immediately following six weeks of instruction, 02, and again two weeks after instruction, 03. The measures at 02 and 03 included both timed variables and power test variables. The data from these measures were analyzed using multivariate analysis of covariance. Based upon the results of the analysis, the following were concluded.

1. While learning profiles reflected opportunity to learn, the instructional sequence SA produced generally lower performance than either of the other two groups; the tendency of all groups was to learn addition first.
2. The amount of learning and retention measured by the timed variables was significantly greater for group AS than for either of the other two groups.
3. At 03, except for addition facts and addition without regrouping, the overall performance of group 1 for the timed variables was significantly greater than that of group SA.
4. At 02 the power test performance of group AS for subtraction was more similar to their performance for addition than was true of the other two groups. At 03, however, the power test performance of group 1 was the most uniform.
5. Differences in the rate of forgetting associated with the three sequences were found only for the power test measures of differences between addition and subtraction were not regarded as important.

6. Instruction with the addition and subtraction algorithms did not result in increases in performance on addition and subtraction number facts. Children's responses for these variables at 02 and 03 showed clear evidence of algorithmic interference.

The implication of these findings was that sequence I was not as productive as sequence AS, but was somewhat more productive than sequence SA. Therefore, this study provides evidence counter to that of other studies which indicate that integrating tasks for instruction is more effective. Finally, this study relates to the sequencing of behaviors for instruction and the learning of complex procedural chains.

Chapter I

INTRODUCTION

The purpose of this study was to investigate the differential effects on achievement of three instructional sequences for the learning of the addition and subtraction algorithms for two-digit numbers. The three sequences were chosen from the same behavioral hierarchy in such a manner that one presented the addition algorithm first, another presented the subtraction algorithm first, and the third presented the two algorithms concurrently. Four sets of measures were used to determine initial performance and the effects of the three sequences on the rate of acquisition of the algorithms, on the achievement level following instruction, and on the retention. These measures provided the basis for comparisons of the relative effectiveness of the three sequences.

Before reviewing research related to this study, it is necessary to describe algorithms so as to differentiate them from other parts of the mathematics content in the elementary school curriculum. Furthermore, while Heimer (1969) has generally defined instructional sequence as, "the order in which the learner interacts with units of content. (p. 494)", the units of content may vary from simple concepts (e.g., Evans, Glaser, & Homme, 1962) to lessons (e.g. Gray, 1970). Therefore, it is also necessary to identify what is meant

by sequence in this study.

The Nature of an Algorithm

The learning of computational algorithms forms a large part of the mathematical experience of children in the present school mathematics curriculum. Apart from procedural ways of determining the usual addition and subtraction facts, the first computational algorithms that are taught and mastered are the addition and subtraction algorithms for whole numbers. Research has indicated that children experience more difficulty in learning a subtraction algorithm than they do in learning an addition algorithm and that neither is readily mastered by some groups of children (Romberg and Planert, 1970; Wiles, Romberg, and Moser, 1972). These findings, of course, validate the personal experience of many elementary school teachers.

The difficulty in obtaining satisfactory learning with these algorithms may be related to the nature of an algorithm, or the previous experiences of the learner. Many of the early experiences of the child have been involved with learning mathematical concepts, but formal algorithms are a type of task that the child has had little, if any, experience in mastering. We can say a concept has been learned when certain combinations of stimuli elicit a consistent response. When an algorithm is learned, however, the appropriate stimuli elicit a chain of behaviors that follows a predictable pattern.

An algorithm may be characterized in terms of instructions or behaviors. An algorithm may be characterized as a termininating sequence of instructions (e.g., Suppes, Jerman, & Brian, 1968). Since it is not possible to observe whether a child has mentally stored instructions or not, an algorithm is characterized here as a sequence of behaviors that can be predicted from the intial stimulus, and that is terminated with an acceptable final response. The subject is presented with a symbolic stimulus, his initial symbolic response together with the original stimulus forms the stimulus for the next response, and so on in a way that may be quite mechanical. This sequence of behaviors will be referred to as a *procedural chain*. The following three characteristics then distinguish algorithmic learning from other learning tasks.

1) There is a sequence of two or more behaviors.
2) The succession of behaviors can be predicted from the original stimuli.
3) The terminal behavior results in an acceptable total response.

For example: if the presented stimulus is $\begin{array}{r} 9 \\ \underline{+6} \end{array}$, and the immediate response is "15", then the learner has not responded algorithmically as only a single response was made. However, if the stimulus is $\begin{array}{r} 29 \\ \underline{+36} \end{array}$, the first response may be to write "5" under the 9 and 6, and "1" above the 2 and 3. The resulting display then becomes

the stimulus for the next response, "6", written under the 2 and 3 leaving the completed example as follows :

$$\begin{array}{r} {}^{1}24 \\ +36 \\ \hline 65 \end{array}$$

While this is not the only possible chain of behaviors that could produce the correct sum, it does demonstrate the three characteristics identified above.

The mechanical nature of "using an algorithm" is also evident from the above example. Hopefully, the mechanics of the entire procedure would be accompanied by the intangible "understanding." This is not, however, necessarily so. A single observed response, particularly if the subject's behavior is not observed as it occurs, may not be sufficient to infer algorithmic behavior. However, additional information as to the prior experience of the subject and the complexity of the stimulus may support such an inference.

The difficulties associated with learning the addition and subtraction algorithms provided the initial motivation for the research being reported, while instructional sequence theory provided a framework in which to examine the learning of these algorithms. Several studies have focused on sequencing strategies for teaching single response tasks, but very little work has been identified that relates instructional sequence and procedural chains. In this regard Short and Haughey (1967) note, "The best strategy for teaching a multiple

discrimination is probably different from the best strategy for teaching a procedural chain (p. 317)."

The Role of Instructional Sequence

Studies that deal with questions of sequence can be classified as feasibility studies and comparative studies. Feasibility studies attempt to show that a developmental system produces an instructional sequence that is in some sense successful. Comparative studies compare the effects of using two or more sequencies that differ in specified ways. The sequences of this study were derived using two sequencing strategies, and the research is basically a comparative study of the effects of using these sequences with implications for the sequencing strategies used to form them.

The systematic development of curricula, such as reported by Kersh (1967), Heimer (1969), and Romberg and Harvey (1969) have made it necessary to examine new questions relating to instructional sequence. The precise relationship of the study being reported to such questions is the content of chapters IV and V, but the general relationship of this study to instructional sequence is as follows:

The instructional sequences here referred to are sequences of behavioral objectives. The research being reported deals with the relative effectiveness of different instructional sequences. In particular, the research focuses upon the relative effects of teaching two procedural chains, addition and subtraction algorithms for two-

digit whole numbers, in three instructional sequences. The learners, of course, must interact with ordered sets of instructional materials intended to facilitate the acquisition of the behaviors, but the order of the instructional materials is determined by the order of the sequences of objectives.

The sequences of this study were constructed from a given hierarchy of behavioral objectives. Heimer (1969) identified four issues related to the use of learning hierarchies in the construction of instructional sequences. They are:

1) How is a learning hierarchy constructed and under what conditions is one considered valid?
2) What is the relationship between an hypothesized learning hierarchy and the associated presentation sequence for instruction?
3) What are the "intellectual skills" that make up a learning hierarchy?
4) What is the connection, if any, between the (logical) structure of the content and the design of the associated learning hierarchy (p. 496)?

The three sequences of this study were constructed from the same behavioral hierarchy and were intended for initial learning, consequently this study relates to the second issue identified by Heimer.

Two of the instructional sequences can be classified as single

task sequences, while the third is an integrated instructional sequence and, therefore, a multiple task sequence. The rationale for these sequences and a detailed description of them is presented later.

Before an elaboration is attempted, however, it is necessary to review the general state of research knowledge with respect to both the algorithms and instructional sequences. A general survey of work done with respect to the addition and subtraction algorithms and with instructional sequences is followed by definitions and further explication of the problem.

Chapter II

RESEARCH RELATING TO THE ALGORITHMS

Studies focusing on the addition and subtraction algorithms can be classified as those that deal with algorithm selection and those that deal with the methodology of teaching the algorithm selected. These classifications are considered in that order followed by a summary of the dependent variables examined in studies relating to the algorithms.

The Selection of an Algorithm

Most of the research dealing with algorithm selection concerns the subtraction algorithm, though some attention has been given to the addition algorithm. The question with respect to the addition algorithm has been, "Is it better to add up or down?", where "better" is defined in terms of speed and accuracy. Though several studies relate to this question, Suppes, Jerman and Brian (1968) credit Murray (1939) with the critical study. Findings prior to Murray's were confounded by problems of sample selection and initial learning experiences. Suppes et al. (1968) report the general conclusion of this matter as "downward addition has been found more accurate while upward addition is more rapid (p. 164)." Subtraction algorithms have received far more attention than have addition algorithms.

The effects of teaching children to use a written "crutch" to record the regrouping when it is required in the usual subtraction

algorithm has been examined by Brownell (1939). While the question of teaching a "crutch" relates strongly to method as well as to algorithm choice, the inclusion or exclusion of such bookkeeping steps results in different algorithms. Brownell's findings indicate that the use of the "crutch" facilitates both the rate of initial learning and the ability to retain this learning. The findings also indicate that children used the "crutch" to enable them to solve new examples and to re-establish proficiency after periods of disuse.

A broader study relating to the selection of a subtraction algorithm was carried out by Brownell and Moser (1949). While the findings of this study also have perhaps more importance for instructional method, this study is the clearest comparison of the results of teaching the subtraction algorithms characterized as the "equal-additions algorithm" and the "subtractive-decomposition algorithm" respectively. Studies prior to Brownell and Moser's that relate to the selection of a subtraction algorithm are confounded by problems of sample selection and initial learning experiences. Suppes et al. (1968) reviewed the entire controversy about which algorithm should be taught, and finally concluded that the Brownell and Moser study avoids the confounding of earlier studies and has practically resolved the issue at present in favor of the "subtractive-decomposition" algorithm that is taught in the schools today.

In summary, the algorithms presently taught in the schools are not challenged by available research evidence, and there is no evidence

to suggest that any extant algorithms should replace them in the curriculum.

The Method of Instruction

Though both algorithms have provided the content for studies of meaningful versus rote methods of instruction (e.g., Overman, 1930), the study of Brownell and Moser (1949), which used only subtraction algorithms, has come to be regarded as the definitive study in the controversy. Brownell and Moser used two different algorithms for whole number subtraction each of which was taught using two methods of instruction: rote and meaningful. Rote meant simply a demonstration of the behavior desired and the insistance that the behavior be practiced until satisfactory performance was obtained. Meaningful instruction was defined by extensive protocols describing in detail the explanations or "rationalizations" that the teacher was to use during instruction. The criterion variables measured were speed, accuracy, retention, transfer, rate of learning, and "understanding." While the findings favored the subtractive-decomposition algorithm as noted earlier, Suppes, Jerman, and Brian (1968), in commenting on the above study stated that, "the most important finding was that it was more economical, in terms of ease of learning and rapid accurate computation, to use rational methods," and "that meaningful teaching leads to greater retention and ability to transfer learning (p. 171-172)."

The controversy over meaningful versus rote method raged for many

years and involved content other than just the addition and subtraction algorithms. While the Brownell-Moser study has provided a consistent bench-mark in the record of that controversy, a summary of the entire issue is provided by Weaver and Suydam (1972). While meaningful instruction as such has not been challenged by research pertaining to the algorithms for the past twenty years, the question of just what is the "most meaningful" method has been raised.

Brownell's study relating to the use of the "crutch" during initial learning sequences (Brownell, 1939) that was discussed in the previous section falls into this category also, as does a more recent study reported by Ekman (1967). Ekman compared the results of using three different approaches to the learning of the addition and subtraction algorithms. The first approach began immediately with a symbolic presentation, the second began with pictures showing the regrouping involved, and a third approach utilized a manipulative aid. In terms of the variables transfer and "understanding", the treatment utilizing a manipulative aid was found to be best; while in terms of computational skill measured immediately after instruction and six and one half weeks later, the first treatment was found to be superior.

The results of Ekman's study point up the fact that the choice of instructional approach is partly based upon the value one places on certain outcomes. However, independent of value choices, research indicates that meaningful instruction during initial learning of the algorithms will give greater transfer of learning, and that emphasis

upon manipulative materials will enhance the meaningfulness of the instruction. If immediate computational proficiency is the primary goal of instruction, rote methods will suffice. The value placed upon certain outcomes has varied with time and investigators, and the dependent variables the investigator utilized reflect this. However some variables have appeared with greater consistency than others.

The Dependent Variables Examined

The dependent variables most frequently examined in research involving algorithms have been the speed and accuracy of computation following the completion of instruction. Often these have been the only variables examined, particularly when the interest was only the relative efficiency of two algorithms or methods of instruction. While no single set of variables has come to be accepted as the norm for research involving the algorithms other variables have come to be regarded as of at least equal importance to speed and accuracy.

When questions concerning instructional method were raised, the set of dependent variables was expanded but not uniformly employed. Brownell (1939) used the variables speed, accuracy, rate of learning, retention, and transfer to items not taught; while Overman (1930) treated transfer as the only variable of interest. Brownell and Moser (1949) examined all of the variables above in addition to measures of the "understanding" possessed by the learner after instruction.

The variable "understanding" is not a well-defined entity but has been operationally defined by many investigators who based their

definitions upon considerations of the algorithms being taught and the instructional experiences provided the children. Osborne (1967), for example, defined a four-dimensional measure of understanding, where the dimensions are "measurable chains" identified as "a) Knowledge-Recall, b) Applications, c) Transfer, and d) Structure-Deduction." A somewhat different approach was taken by Brownell and Moser (1949) in which they coded and categorized the verbal and written responses of children as the children offered their explanations of why a particular example was done correctly.

Aside from the obvious consideration of choosing variables that will answer the questions asked, the above review indicates that a large selection of variables is desirable and that they should include at least measures of speed and accuracy. Measures of the rate of learning, retention, and of some variables associated with understanding are also desirable, but again no one set of variables have come to be regarded as a standard set.

To summarize this chapter, the above review has shown that evidence exists to support the inclusion of the usual addition and subtraction algorithms in the elementary school curriculum. The evidence strongly suggests meaningful methods of instruction be employed as opposed to simply supplying examples and practice. The desirability of a large selection of dependent variables that includes more than just measures of computational efficiency following instruction has been indicated for research dealing with the algorithms by the fact that important

differences have been overlooked when this was the only variable considered. However, apart from a study reviewed in a following section dealing with the use of integrated sequences (Wilcs, Romberg, and Moser 1972), research has not been directed to the order in which the algorithms are taught, or to the effects of integrating instructional sequences designed to teach these algorithms.

Chapter III

RESEARCH WITH INSTRUCTIONAL SEQUENCES

General Classification of Studies

Research studies dealing with instructional sequence may be classified as those which attempt to find an optimal ordering of a set of concepts or skills (e.g., Suppes, Jerman, & Brian, 1968; Jerman, 1970; & Farris, 1971), and those which compare different orderings of instructional elements designed to teach a set of tasks. This second group of studies may be further differentiated into those that deal with randomly ordered sequences and those that deal only with rationally ordered sequences.

Studies in the first group have attempted to find optimal orderings for a set of tasks based upon representational mode (e.g., Farris, 1971) or relative difficulty (e.g., Jerman, 1971). However, there is no evidence that an optimal instructional sequence is one in which the "easier" content is taught first when dependency relationships provide no guidance for sequencing.

Studies dealing with random orderings of instructional stimuli have been motivated by problems that arise when constructing programmed learning materials. King (1970), while cautioning against over-generalization, summarizes this entire line of research as follows:

> "...research suggests that varying highly refined sequences of instructional stimuli does not make much difference in effectiveness of instruction, as long as concept order is preserved (p. 238)."

A great deal of recent research has associated instructional sequence with the general area of discovery learning. Variables such as the timing and amount of verbalization, the ordering of examples and exposition, etc., have been examined by many investigators (for example, Worthen, 1967; Scott, 1970; Rector and Henderson, 1970). As an illustration of the general procedure Nelson and Frayer (1972) varied the order in which the definition of a concept and examples of the concept were presented to the learner. While these studies have provided information about sequences of instructional moves as related to single concepts, the results of these studies do not bear directly on the problem of sequencing several units of content.

To summarize this section, it seems reasonable to expect observable effects on achievement measures resulting from varying the order in which the algorithms are taught. While the algorithms are procedural chains rather than concepts, sequences that vary the order in which the addition and subtraction algorithms are taught do seem to meet King's criteria for finding possible differences due to instructional sequence.

It should also be noted that most, but not all, of the studies reported here employed programmed materials. The use of programmed materials does provide a strong control of teacher variables and insures uniformity of treatment. But, as Worthen (1967) noted, when the generalization is to be made to classroom settings it is not really a viable choice. This is particularly so when generalizations are to be made to activity-based programs that emphasize student interaction and inquiry

and to populations of students that read poorly if at all. Many aspects or programmed instruction can be incorporated in a classroom setting, however, as noted by Kersh (1967);

> "...the principles and techniques which underlie self-instructional programming can be employed equally as well in the development of suitable classroom instructional materials and procedures. The result may be very similar in appearance to classroom procedures which are presently employed by teachers, but the resemblance may end there. There is no greater similarity between conventional classroom techniques and programmed classroom techniques than there is between conventional self-study materials and programmed self-instructional materials (p. 81)."

The Use of Integrated Sequences

The studies reviewed in this section report investigations of the comparative effects of two or more instructional sequences, some of which are identified as integrated in respects that the other sequences are not. The integrated sequences are formed by merging content units that logically may be sequenced consecutively. The reasons for merging the units vary with the study, but generally include such considerations as similarities among the stimuli or responses, the existence of common prerequisites, or relationships among the merged content units that form new concepts or relations. When learning hierarchies have been used as an aid in selecting instructional sequence, the form of the hierarchy has sometimes suggested to writers the integration of disjoint branches to form an integrated instructional sequence.

Newton and Hickey (1965) report an experimental comparison of the

effects on achievement resulting from using twelve different instructional sequences to teach the concept "gross national product" along with associated subconcepts and definitions. The sequences were chosen from an existing set of 59 programmed learning frames on the basis of a logical analysis of the major concept and presented to learners in the form of programmed booklets.

The analysis of the major concept identified three subconcepts, one of which was eliminated from the instructional sequences. The two possible orderings of the remaining subconcepts formed the levels of the first variable.

The second variable, a "position variable", was intended to reflect the timing of the learning of the subconcepts. Basically the concepts were to be learned together and remembered until such time as they were needed in the definition of the main concept, or they were to be learned one at a time as they were needed in defining the concept "gross national product". Practically, this was accomplished by placing a set of frames dealing with other related definitions either before the two subconcepts were taught, after they both were taught, or after the first, but before the second was taught. In general, the sequences were said to be intergrated when the subconcepts were taught consecutively with no intervening frames. A third variable, direction, was operationally defined by sequencing the frames toward the definitions of the concepts (all three of them) or away from the definitions. This variable may be thought of as a "discovery variable."

The twelve sequences then were determined by using all of the possible combinations of the specified levels of the preceding variables. These variables and their levels are as follows:

1. The two orders in which the two subconcepts could be presented (order - two levels),
2. The position in the program of the set of related definitions, i.e., before the first subconcept, after the second subconcept, or between the two subconcepts (position - three levels), and
3. The direction of the program with respect to the definitions of the three concepts, begin with definitions or end with definitions (direction - two levels).

The experimental subjects were college students "unsystematically assigned" to groups. They were allowed 50 minutes to complete the program and take a ten-item multiple-choice test. The criterion variables were: time required to complete the program, number of wrong responses during learning, and the multiple-choice test score.

The results of Hickey and Newton's study are difficult to interpret because there are few significant main effects, or even trends. However, their general conclusions include the following statements. "Performance was faster when principles were stated first and when subconcepts were learned together rather than separately. Ss made significantly poorer scores on a multiple-choice criterion test under conditions where learning of the two subconcepts was most remote from integration of these to form the major concept (Newton & Hickey, 1965, p. 140)."

The integration of the subconcepts in the instructional sequence was thus judged to produce performance sooner, and was associated with higher post-test scores.

Short and Haughey (1967) report a study that compared the effects of using two sequencing strategies called "multiple-concept" and "single-concept" defined as follows:

> "The multiple-concept strategy presents simple descriptions of several related concepts at the beginning of instruction. Increasingly complex material pertaining to all these concepts is gradually introduced. ...The single-concept strategy presents one concept at a time, proceeding from a simple description of the single concept to more complex descriptions of the same concept. After the concept has been presented in all its detail and complexity, a second concept is introduced, described in detail, and then a third concept is introduced, etc. (p. 317)."

The multiple-concept strategy produces an integrated instructional sequence for the concepts involved.

The strategies outlined above were used to generate two instructional sequences designed to teach multiple discrimination tasks. A multiple discrimination is characterized as follows: "In learning a multiple discrimination a student learns to make different responses to similar stimuli that previously evoked an undifferentiated response, (Short & Haughey, 1967, p. 317)." A pilot program in which an artificial multiple discrimination task was taught favored the multiple-concept sequencing strategy. Following the pilot, Short and Haughey selected concepts from two content areas, science and language arts, and constructed two sequences using the different strategies. The science

materials were embodied in a fixed pace, slide-tape presentation, and the language-arts materials were presented in program text form. The materials were presented to randomly selected groups of fifth grade students on five consecutive days, one hour per day. Criterion tests were administered as pretests, post-tests, and one-week retention tests.

The results are summarized as follows: ". . .all comparisons of student performance favored the groups receiving the multiple-concept sequence (Short & Haughey, 1967, p. 318)." However, only the science materials produced differences that were regarded as statistically significant. Interestingly, the authors note that the differences can be accounted for in terms of the necessity of learning a new concept each lesson. In other words, the best learning conditions for these subjects, with these materials, were when examples of each concept were present in each lesson.

Short and Haughey emphasize that the kind of task their study is concerned with is a multiple-discrimination task, and that the "best strategy for teaching a multiple discrimination is probably different from the best strategy for teaching a procedural chain (p. 317)."

Gray (1970) reports a study of the relative effects of teaching an integrated and non-integrated learning sequence on the acquisition and retention of mathematics and science behaviors by fifth-grade children. Three "quantitative science behaviors" were identified as the final objectives of the learning sequence. A task analysis of these three behaviors resulted in a hierarchical arrangement of twenty-five objectives involving both mathematics and science. Two sequences were selected from this hierarchy. One sequence separated all of the

mathematics behaviors from the science behaviors and taught the former first. The other sequence was integrated with respect to these two sets of behaviors so as to emphasize relationships between them. A set of twelve lessons to be taught in twelve classroom sessions were developed for each of the two sequences.

Gray randomly assigned thirty intact classes to the two sets of lessons. The criterion variables were class mean scores on a test of the behaviors to be learned. The test was administered following the completion of the lessons and again nine weeks later.

Gray's summary includes the following statements: "The integrated sequence produced a significantly higher overall performance than the non-integrated sequence in acquisition of the mathematics behaviors although there were no significant differences in the effects of the sequence on the rate of forgetting. The two treatments had no differential effects on the overall performance or the rate of forgetting with regard to the science behaviors." Gray further concludes that the integrated sequence is "generally superior to the non integrated sequence in facilitating acquisition of the mathematical behaviors for the population defined in this study (p. 3 of the abstract)."

The three studies reviewed above are unanimous in their conclusions that an integrated sequence proved superior to a non-integrated sequence. However, some qualifications of the findings need to be noted.

1. In all of the studies reviewed the subjects were the ages of fifth grade students or older; results obtained with

younger children might be different.

2. With the possible exception of a few of Gray's tasks, the content to be learned is a concept or a set of concepts. If the content material were procedural chains, the results obtained might be different.

An Earlier Study

Wiles, Romberg, and Moser (1972) report a study that is directly related to this research. Two sequences of behaviors associated with the addition and subtraction algorithms for whole numbers were constructed. In one of the sequences (I) the behaviors were ordered so that both algorithms were presented concurrently. In the other sequence (S) the behaviors associated with the addition algorithm were presented first. Instructional activities embodying the two sequences were then developed. Children of two second grade classes were randomly reassigned to two groups, one of which was to use the S activities and the other the I activities.

The sequence S followed the usual sequencing of these algorithms for initial instruction, addition followed by subtraction. Romberg and Planert (1970) found that children experienced much more difficulty in learning the subtraction algorithm than they did in learning the addition algorithm when they were presented in the usual order. In fact, it was the general dissatisfaction with this outcome that led to the hypothesis that overall performance, and performance with the subtraction algorithm in particular, could be improved by the use of an integrated sequence of

the behaviors associated with the two algorithms. Wiles et al. was one attempt to evaluate this hypothesis.

The rationale for the sequence I was as follows:

1. Instruction to this point in the experience of the children emphasized the simultaneous introduction and development of addition and subtraction.
2. The regrouping associated with the addition algorithm as "carrying" is the reverse of the regrouping associated with the subtraction algorithm as "borrowing".
3. Addition and subtraction of whole numbers are both counting processes.
4. Addition and subtraction of whole numbers are inverse mathematical operations.
5. The common mechanical characteristics of "begin at the right", and vertical form employed suggest a parallel development of the algorithms (p. 4).

Performance of the groups was measured on the eighteenth day of instruction by a 20-item test. Growth and achievement levels during the course of instruction were estimated by the use of item-sampled test forms administered every third instruction day. While only a few of the tests of group differences were judged to be statistically significant, and these were all differences on the item-sampled tests, the direction of every reasonable comparison of all test scores favored the group experiencing sequence S.

This outcome is clearly contrary to what would be expected from considering the previously reviewed research. However, the superiority of sequence S is not clearly established for the following reasons:

1. The sequence I may not represent the strongest possible emphasis upon the similarities between the algorithms. The integration of the behaviors was accomplished by grouping them in levels and then writing activities that encompassed the entire group. However, the activities did not specifically call the attention of the learner to the similarities and differences between the behaviors. To do so, of course, introduces instructional material into the sequence that is not present in sequence S and so some control for opportunity to learn would be lost. Logically, however, such a procedure should provide a stronger I sequence.
2. The selection of dependent variables was minimal. Performance data prior to instruction were not available, and the eighteenth-day test measured only the ability of the students to produce correct sums and differences. While the item-sampled test did contain "understanding" items, it did not contain any scales apart from a general addition and a general subtraction scale. Furthermore, no measures of retention were made.
3. A satisfactory level of performance with the subtraction algorithm was not achieved by either group in the 20-day

treatment period. Had the treatments continued until some higher level of performance was reached, the results would be of greater interest.

4. A feasibility study of the I treatment and the profile-generating tests reported by Romberg and Wiles (1972), a group of students had an achievement profile that more nearly met the desired expectations of the I treatment in that they achieved generally higher mastery levels than groups I and S of Wiles et al. (1972). However the group of the feasibility study is not comparable to the groups of the comparative study in that it was a selected high-ability group. While Cronbach and Snow (1969) indicate that established ability by treatment interactions have been few in number, the possibility of such an effect is evidenced by the apparent success of this high-ability group with the sequence I.

In spite of the factors noted above, a final summary of this earlier study must be that the most reasonable hypothesis for the observed results is that the attempted simultaneous learning of two somewhat similar procedural chains resulted in interference between the two.

The relationship of the research reviewed in this chapter to the study being reported involves a systematic plan of curriculum development. The following chapter, Chapter IV, describes the evolution of systematic ways of constructing curricula and instructional sequences. Chapter V then outlines the specific plan of curriculum development

which forms the context of this study, and specifies the relationship of the above research to the problem of this study.

Chapter IV

SYSTEMS FOR CONSTRUCTING INSTRUCTIONAL SEQUENCES

Definition of Sequence

The core of the difficulty in constructing an instructional sequence is the identification of the elements of content and the determination of the order in which the child is to experience them. Heimer (1969) in his review of sequence theory development states: "Instructional sequence means the order in which the learner interacts with units (elements) of content (p. 494)." Accordingly, an instructional sequence is defined as the elements of content and the order in which the learner is to interact with them. What the elements are depends upon the manner in which they are developed from the given content, they may be frames, concepts, tasks, or behaviors depending upon the characteristics of the system used to generate them.

Various systems have been devised to provide some guidance for the selection and ordering of the elements, but these systems function as techniques to guide art rather than as algorithms that determine an outcome. In regard to this general problem Heimer notes: "...at present there are few precise, coherent, and empirically testable sets of hypotheses for identifying and guiding those instructional decisions. In short, there are no adequate teaching algorithms (p. 494)."

The systems approach to curriculum construction at present involves much more than just instructional sequence. While it is not the purpose

of this paper to deal with curriculum construction in general, some discussion of the application of systems approaches to the entire enterprize is necessary to clarify the role of sequence construction within it. After a brief review of systems applications to curricula, the development of systems for deciding instructional sequence will be discussed.

It may be argued that the construction of curricula has always been done in a systematic fashion, but what is in reference here is the deliberate application of the principles and practices of systems analysis to curriculum development. This application has of course required the development of new concepts and new procedures. As systems concepts and procedures have been developed in the context of curriculum construction, those particular aspects that deal with instructional sequence have developed also. The development of these two aspects have been at times independent, with the systems for sequence construction being absorbed as subsystems of the larger system of curriculum construction.

Systems Approaches to Curriculum Construction

Silvern (1965) traces the development of systems approaches in education from Herbart's "formal steps" in the early nineteenth century in what is clearly a context of curriculum construction. Silvern continues the trace through industrial education, engineering education, vocational education, and military applications through World War II. In his review Silvern then identified 46 "better-known (system)

conceptualizations subsequent to World War II (p. 41)" in the area of education and training.

The 46 conceptualizations are characterized in terms of five factors, "(1) nature of the model, (2) use of feedback to provide system stability, (3) use of simulation, (4) use of analysis, and (5) use of synthesis (pages 41, 42)." While all of the conceptualizations discussed involve curriculum many of them are primarily concerned with the construction of curriculum. The use of teaching machines, programmed instruction, relatively inexpensive new media and computer technology are evident as increasingly important factors in the conceptualizations cited.

Baker and Schutz (1971) as editors of a book published six years after Silvern's extensive report state in their overview of the applications of systems to instructional product development, "Programmatic instructional development efforts conducted during the last few years are only just beginning to produce generalizable procedures and strategies (page xv)." While no review of what has been done is attempted by Baker and Schutz, their intent as editors is to present an outline of "common systems' characteristics," that, "have proved useful in the programmatic development of instructional products (p. xv)."

A chapter of Baker and Schutz written by Popham and Baker (1971) and titled "Rules for Development of Instructional Products" identifies a product development cycle of seven stages. These stages are: formulation, instructional specifications, item tryout, product development, product tryout, product revision, and operations analysis. Since

it is not the purpose of this present report to deal extensively with the general area of curriculum development these seven stages will not be detailed here. The interested reader is referred to this chapter. For present purposes it is sufficient to note that decisions about instructional sequence, as defined earlier, are made at the second stage of development. But in keeping with the systems concept of corrective feedback, the results of the operations of later stages may require some adjustments in instructional sequence.

Systems Approaches to Instructional Sequence

The various systems that have been devised for constructing instructional sequences follow a general developmental trend. The initial focus was on constructing and ordering frames for a programmed learning experience that was to teach a single concept or principle (Evans, Homme, & Glaser, 1962). Hickey and Newton (1964) expanded this focus to include several concepts, and Kersh (1967) describes a system that involves both concepts and the associated behaviors that the student is expected to evidence. Furthermore the system described by Kersh is intended to have applicability to the general classroom as earlier noted (see Chapter III).

The literature dealing with systems of sequence construction is descriptive in nature; studies of comparisons between different systems of sequencing are not available. Descriptions of the systems frequently occur in explanations of how two or more sequences were constructed where the system contained no rules for choosing between the alternatives.

The procedure followed in this review is to briefly survey systems of sequence construction in the order of their development and identify the common characteristics.

Evans et al. (1962), reports a system called RULEG that was developed as an aid for the construction of programmed verbal learning sequences. The elements of this system are rule statements and examples of rule statements. For instance, the distributive law for whole numbers could be a rule, and $2(3 + 5) = 2(3) + 2(5)$ an example of the rule. The rules and examples are formed into matrices that indicate reltionships among them. The matrices are only used to help organize thinking, and with reference to these organizing matrices, frames are written and a sequence of frames is chosen.

Hickey and Newton (1964) describe a "knowledge space" which they regarded as a larger frame of reference than RULEG. The methods of RULEG were referred to as most helpful "in the formulation of brief subsequences," but provided "no clear-cut formulations for generating the overall program organization, (Newton & Hickey, 1965, p. 140)." Their expansion of RULEG called "knowledge space" is operationally defined by a tree-graph display of concepts, subconcepts, and attributes. These conceptual elements are the elements of content that are ordered to form the instructional sequences. An ordered set of frames provides the final instructional form of a sequence.

Gagne (1968) and others have applied the methods of task analysis to problems of designing instruction. The elements are behavioral objectives arranged in hierarchies that provide a partial ordering of the

elements. This partial ordering of elements is a characteristic shared with the tree-graph displays of "knowledge space" that greatly facilitates the selection of sequences of the elements. Walbesser and Eisenberg (1972) provide a review of theoretical positions and research relating to the use of such hierarchies.

The embodiment of a sequence of behaviors may also be an ordered set of frames, but other instructional forms are possible. Gray (1970), for example, used the methods of task analysis to aid him in selecting sequences of behaviors, but the embodiments of the sequences were sets of lessons. Each lesson was designed for one instructional period in a self-contained, traditionally organized classroom.

Kersh (1967) outlines a system of curriculum development called TRAC that contains provisions for selecting and ordering elements of an instructional sequence. The component of the total system that deals with instructional sequence involves two behavioral hierarchies, one based upon considerations of what the learner must do to demonstrate that he knows some particular concept, relation, etc., and the other based upon considerations of what the learner "must do in order to acquire both the knowledge and the complex behavior (p. 85)." The first hierarchy seems to be concerned with products while the second is concerned with processes. A general assumption is that a teacher will be required if the behaviors of the second hierarchy are to be achieved. These two hierarchies are constructed and integrated to form sequences of behaviors. As noted earlier, the instructional form may be a set of frames or a set of lessons.

While the primary concern of this review has been that part of a total system that deals with instructional sequence, the fact that there is a total system viewpoint which takes specific notice of instructional sequence represents an advance in the art of sequence development. The total curriculum viewpoint of TRAC is seen in the following quote from Kersh (1967).

> "It should be possible for an expert in human learning and a subject matter specialist to prepare, in advance, an outline of the learning process, just as an engineer does in designing, on paper, the structure and systems he builds. Then, from these 'instructional designs', it should be possible for programmers and materials development specialists in our schools and colleges to actually 'build' the instructional systems, try them out, and, if necessary, send them 'back to the drawing board' to be modified (p. 82, 83)."

Summary

Systems for constructing instructional sequences have evolved from those that were concerned only with the orderly arrangement of concept rules and examples to those that are components of a complete curriculum development project. However, every system for constructing an instructional sequence must provide for analysis of the given content into elements that can be linearly ordered to form a sequence. The most comprehensive systems involve the use of behavioral hierarchies constructed by the methods of task analysis, though the knowledge space idea of Hickey and Newton (1964) has not been discredited by research.

The relative effectiveness of the systems discussed has never been assessed; this is undoubtedly due to the essentially artistic nature of

the endeavor. As noted earlier, algorithms do not exist for deriving either the elements of the sequence or the order in which the elements are to occur, but evolving techniques can be identified in which behavioral hierarchies have become increasingly important. Though some suggestions for assessing the validity of a behavioral hierarchy have been made (see Eisenberg and Walbesser, 1971), extensive reports of "valid hierarchies" are not generally available.

The form of the instructional materials embodying an instructional sequence varies widely enough to include both frames for programmed learning sequences and lessons for use in a conventional classroom setting. While the task of determining sequence is separate conceptually from the instructional form, the eventual embodiment of the sequence is a constraint to be considered when choosing a sequence of content elements.

Chapter V

THE PROBLEM AND THE SEQUENCE SELECTION MODEL USED

The problem was investigated within the scope of the research activities of the Analysis of Mathematics Instruction Component of the Wisconsin Research and Development Center for Cognitive Learning (AMI). The AMI project staff is employing a systems approach in developing a new elementary school mathematics program, Developing Mathematical Processes (DMP) (see Harvey, Romberg, & Fletcher, 1969). Though the developmental plan as outlined by Harvey, et al. (1969) consists of the five phases: definition, analysis, pilot development, small scale field test, and large scale field test, only the analysis phase is of interest here. This chapter consists of a discussion of the analysis phase followed by a general description of the type of instructional activity employed, some definitions, and a statement of the problem in terms of those definitions.

The Analysis Phase

The analysis phase begins with the identification of the content to be taught and an analysis of this content into a hierarchy of behavioral objectives using the techniques of task analysis (Gagne, 1968). The hierarchy is the focus of an additional instructional analysis that partitions the hierarchy into sub-units called topics, orders the behaviors within the topics, and orders the topics to form an instructional

sequence of the behaviors of the entire hierarchy. In general the instructional sequence begins at the lowest levels and proceeds to the top through the various levels in a manner that is consistent with the partial ordering present in the hierarchy.

In the analysis phase of the developmental plan sequences of behaviors are selected. After these sequences are chosen, instructional activities are written that should enable the learner to master the behaviors. The set of student activities finally produced could be designated as the instructional sequence. In fact, changes in the original sequence of behaviors can be made of the basis of information provided by any of the later developmental phases. However, since the initial decisions involve choosing ordered sets of behaviors and since the content of the final product is least ambiguously identified in terms of behavioral objectives, in this study ordered sets of behaviors are identified as instructional sequences.

The entire analysis phase of the developmental plan, then, begins with the identification of the content to be taught and ends with the production of a tentative topic outline. A crucial part of this component is the behavioral hierarchy and the instructional sequence derived from it.

As has already been shown in the previous chapter, the use of hierarchies of behaviors as the basis of an instructional sequence is not a new idea. In fact Heimer (1969) was responding to the widespread use of such hierarchies when he identified the four issues quoted

in the introduction of this report. The study reported here, as was earlier noted, bears on the second issue, "the relationship between a hypothesized learning hierarchy and the associated presentation sequence for instruction (p. 496)." The relationship between this issue and the study will be discussed after a description of the instructional activities employed and the presentation of a set of definitions that help clarify this relationship.

Instructional Activities

The sequences of this study were embodied in activities that usually involve several behaviors. The grouping of behaviors in a particular activity allows their designation as preparatory, mastery, or review with respect to a given activity. While behaviors are probably mastered one at a time, progression toward mastery does not seem to be a behavior-by-behavior procedure. An activity may be oriented toward discovery, exposition, or at some point between these two extremes; it may be designated for a single child or for groups of children, but in general it will specify what a child is to do (by himself or in a group). A rationale for this particular instructional form is given by Romberg, Fletcher and Scott (1968) where acceptable activities are described as follows.

> It will be necessary to develop instructional tasks (activities) to introduce each basic unit of the program that meet generally accepted criteria regarding motivation. Three criteria for these tasks are:

a) that they come from the child's environment

b) that the child sees them as problems to which he is able to find solutions, and

c) that the tasks provide immediate and informative feedback (p. 4).

Also,

> Motivation is further maintained by relying upon functional transition from state to state. ...In essence, then, transition will be preceded by a practical problem or a functional reason for having to learn something new. ...These reasons will always be made explicit and the child will see the functional basis for having to acquire a new skill. In no case will this require the child to "learn" by rote without providing the practical reason for such learning (p. 4).

For further specification the interested reader is referred to the paper by Romberg, et al, (1968) or to the appendix of this present report that contains activities used in this study.

Because of the difficulty in controlling teaching behaviors, many investigators have used programmed materials when exploring questions of instructional sequence. However, the use of more conventional materials in research involving such questions is not without precedent (e.g., Gray, 1970; Worthen, 1967). For young children who read poorly if at all, the activity format used seems more appropriate than programmed materials while possessing adequate structure to control variance due to teacher behvaiors. This is believed to be particularly so when teachers are familiar with the instructional approaches characteristic of DMP.

The implementation of DMP normally involves modification, extension, or even replacement of the given activities by the teacher in ways which are consistent with the activity approach. However, for this study, the activities were specified and interpretations were standardized to ensure that the children did the intended activities.

Designing and constructing instructional activities according to a specified sequence of behaviors is not a trivial matter. The characteristics of the chosen instructional format may dictate some alterations in the instructional sequence selected, however the sequencing questions of this study arise from the necessity of choosing ordered sets of behaviors from a hierarchy and of ordering these sets to form an instructional sequence. The questions of interest involving sequences will be made explicit using the following terms.

Definitions

1. The behaviors of a hierarchy are partially ordered by the relationship precedes. One behavior is said to precede another if the performance of the first is judged to be necessary for the performance of the second. If one behavior precedes another, the first will be called prerequisite to the second.
2. A task is one or more behaviors. In general, if a task contains two or more behaviors chosen from a hierarchy, they are linearly ordered in a manner that does not violate the partial ordering of the hierarchy.

3. A behavior precedes a task if the behavior is not in the task but does precede some behavior in the task. A behavior that precedes a task will also be said to be prerequisite for the task.
4. One task precedes another if at least one behavior of the first task precedes the second task. A task that precedes another will also be called subordinate to the second. (While it is possible that two tasks chosen from the same hierarchy would be mutually subordinate, the process of selecting tasks would normally continue until this is not true.)
5. Two tasks from the same hierarchy are called coordinate if neither precedes the other.
6. When coordinate tasks are linearly ordered the resulting sequence is called a single task sequence.
7. If two or more coordinate tasks are merged in some way to form a new task, the newly formed task is called integrated with respect to the original tasks.
8. The ordering of all the tasks selected forms an instructional sequence for the hierarchy from which the tasks were chosen.

Tasks are ordered to form an instructional sequence in the following manner. When tasks are subordinate the sequence is clearly indicated; however when tasks are coordinate, the sequence is not determined. In addition to the possible permutations of the coordinate tasks, two or

more of them could be integrated to form a new task. When the behaviors of coordinate tasks have similarities in stimulus or response conditions, or are related in the logical structure of the content in ways not indicated by the hierarchy, the possibility of integrating the tasks becomes particularly attractive.

The important observation is that in any non-trivial hierarchy of behaviors, pairs of coordinate tasks can be identified and some sequence involving the behaviors of these tasks must be selected. The final selection of an instructional sequence may be guided by philosophy, psychology, or by further examination of the structure of the discipline. A taxonomy of behaviors may also help decide the issue, but the decision of which sequence to choose is made on considerations additional to the task analysis.

This study provides some knowledge of the consequences of integrating two closely related coordinate sequences to form an instructional sequence; the study is in this way related to Helmer's second issue.

The Problem

The purpose of this study was to investigate the differential effects on achievement of the three instructional sequences AS, I, and SA; each of which was designed to teach the usual addition and subtraction algorithms for two-digit whole numbers.

While this study is related to Wiles, Romberg, and Moser (1972), in that earlier study the effects of the sequence, subtraction followed

by addition, were not investigated. Furthermore, a more detailed examination of effects is provided by this study as specified in Chapter VI.

In terms of the preceding definitions, the primary problem of this study was the comparison of the effects of using three instructional sequences on various performance measures. The sequences are the two single task sequences determined by the permutations of two coordinate tasks, and a sequence formed by an integration of the two coordinate tasks. One of these coordinate tasks (A) is the set of behaviors associated with learning the addition algorithms for two-digit whole numbers. The other task (S) is the set of behaviors associated with the subtraction algorithm. These tasks were chosen from the behavioral hierarchy presented in Appendix A, and are precisely identified in that appendix. The integrated sequence formed from A and S is called I, the other sequences are called AS (for addition followed by subtraction) and SA (for subtraction followed by addition). These instructional sequences were embodied in activities typical of DMP as described earlier. The labels for the sets of activities are the same as those for the sequences. Examples of the actual activities used are found in Appendix B. The comparisons among the sequences were made to provide information relating to the following specific questions.

I. What differences attributable to instructional sequence can be found on measures of the following variables?

1. the rate at which the algorithms are learned,

2. the achievement levels for the terminal behaviors observed after instruction, and

3. the retention of the terminal behaviors.

II. Do measures of proficiency with the addition and subtraction facts show significant changes due to the incidental drill inherent in learning the usual algorithms?

III. Are there any interactions of instructional sequence with IQ levels on any of the above variables?

IV. To what extent is the learning hierarchy valid as measured by the methods suggested by Eisenberg and Walbesser (1971)?

Of particular importance to mathematics education are the following concerns. While there is evidence (Wiles et al., 1972) that the subtraction algorithm is more difficult to learn than the addition algorithm, it was not known how this might be related to the order in which they are presented, nor if the most efficient instructional sequence is to teach the "easier" task first. Additionally, while Wiles et al. indicated a definite trend favoring the single task sequence (addition then subtraction), the factors cited earlier in the review of that study left sufficient doubt to justify the closer examination of the relative effects of using an integrated sequence. The importance of these two algorithms in the elementary mathematics curriculum alone justifies the search for an optimal instructional sequence. This search is further justified by the large proportion of the total mathematics instruction time devoted to learning algorithms in general.

The sequences used in this study are related in definite ways to a behavioral hierarchy and since such hierarchies are integral components of recent mathematics curriculum development projects (e.g., Kersh, 1967; Romberg and Harvey, 1969), this study has still further importance for mathematics education. With respect to learning hierarchies, two general questions that extend beyond mathematics education can also be identified:

1. How does the use of an integrated instructional sequence affect learning compared to single task instructional sequences of the same tasks, when the content is identified as closely related procedural chains, and the subjects are elementary school children in the age range 7-9 years?

2. What effects on learning can be attributed to different orderings of coordinate tasks when these tasks are procedural chains?

While the research reported does not completely answer these questions, it does provide new information.

Chapter VI

DESIGN AND PROCEDURES

Introduction

The experimental design and procedures that governed the conduct of the study form the content of this chapter. The implementation of the design and procedures form the content of the following chapter.

The experimental design includes random assignment of subjects to three treatment groups. The treatments are the three sets of instructional activities that embody the sequences AS, SA, and I; all of these treatments provide instruction in the usual algorithms for the addition and subtraction of two-digit whole numbers. No control group is included in the design since algorithms are not likely to be learned independent of instruction.

The design includes several measures of student achievement. These are taken prior to instruction, during instruction, immediately following instruction, and two weeks after the termination of instruction. The mean scores of the groups provide the basic data that bear upon the questions of the study.

The experimental design is summarized in Figure 1. This design is explicated by discussions of the subjects, treatments, observational instruments, and analysis in that order. In this discussion and in the following chapters the labels AS, SA, and I are used for

the sequences, the treatments, and the experimental groups. This labeling should cause no confusion since the context of the discussion in which these labels occur is clearly one of sequence, treatment, or group. The symbols 01, P, 02, 03 refer to observations of group performance.

Group	Assignment			Treatment			
AS	R	01	P	AS	02	---	03
SA	R	01	P	SA	02	---	03
I	R	01	P	I	02	---	03

Figure 1: Experimental Design

Subjects

The selection of Ss is accomplished within several constraints. Some of these constraints relate to the specification of population and others are imposed by the research paradigm.

The instructional treatments must be in the activity format discussed in earlier chapters so that inferences may be made to instructional approaches typical of DMP materials. The effective use of such approaches however, requires experience in using them by both teachers and students. Therefore, extensive experience with DMP on the part of both teachers and children is a primary constraint.

Within this population the specification of entering behaviors is also necessary. Basically the necessary prerequisites include the ability to form sums and differences commonly referred to as the basic

facts, and the inability to perform calculations using the usual algorithms. Instant recall of a sum or difference is not required, but rather the ability to form a sum or difference in a reasonable time. The precise specification of these behaviors is found in Appendix A. The intent of this specification of behaviors is to require that a suitable subject possesses the necessary prerequisite behaviors but does not possess the behaviors to be learned.

The prior use of DMP materials up to the point of beginning the experimental treatment insures that students have the prerequisite behaviors and are familiar with the concepts of addition and subtraction used in all of the treatments. To insure that Ss do not have the behaviors to be learned however, a pretest that is described later is used.

The constraint of involvement with DMP restricted the number of schools that could supply subjects for this study. While the population of students that could potentially experience DMP materials is as large as might be desired, the actual set of schools using these materials at the time of the experiment was small. This is true since these materials were in the development stage and were not generally available at the time of this study.

In addition to the constraints on previous experience, the research paradigm requires the random assignment of subjects to three treatment groups. The Ss were ranked by IQ, grouped into consecutive

blocks of three, and one member of each block of three randomly assigned to each treatment. To insure approximately equal numbers of boys and girls in each group, this procedure was carried out separately for boys and girls.

The presence of all the subjects in a single school allows for the possibility of a strong control for history during treatment. The corresponding constraint imposed by the paradigm is that each group study arithmetic the same days at the same hour of the day, and that pre- and posttesting be accomplished in large group settings.

Treatments

The treatments are the sets of activities embodying the previously identified instructional sequences. The following sections describe the activities and their development, and discuss the pacing and management of instruction. Examples of activities used are found in Appendix B.

The Construction of the Activities

To insure that the activities used did not vary across treatments, the activities for Treatment I were written first, and the activities for AS and SA were derived from them. This was accomplished by separating each activity of I into an addition component and a subtraction component and reordering these components according to the previously identified requirements of the sequences AS and SA.

However, some departures from a strict partitioning were made.

The activities for Treatment I were written so as to encompass the behaviors of Tasks A and S that appear on the same level of the task analysis in Appendix A. Furthermore, several activities direct the children's attention to similarities and differences among these behaviors that are not indicated by the task analysis. When the activities of I were partitioned, any aspects of an activity that deal explicitly with these similarities and differences were omitted and do not appear in the activities of AS or SA.

Another departure from the strict partitioning and resequencing of the partitions of the I activities according to the requirements of AS and SA, is that during the course of the subtraction activities of Treatment AS an occasional addition item is included, and similarly, for Treatment SA an occasional subtraction item is included in the addition activities. These items are included solely as skill maintenance items. The teacher may acknowledge the existence of such items, but no specific teaching toward the item is to be given at that time. The student is told simply to work all items in the context of the given activity.

The activities follow a progression in the modes of the stimuli presented and the responses required. Initial activities employ physical stimuli and require physical manipulations with objects to determine the sum or difference, which is then to be recorded symbolically in the usual notation. In the activities that follow,

the stimuli gradually change from a physical to a symbolic form while the physical manipulations are represented by using expanded notation, bookkeeping steps, and finally the usual compact notation algorithms. The model for the expanded notation symbolism is the grouping and regrouping of objects, and models for the compact notation symbolism are the expanded notation symbolism as well as the grouping and regrouping of objects. Using this procedure the learner should always be able to model the slightly more abstract task with a previously learned, less abstract model.

Pacing of Instructional Activities

The management decisions regulating the pacing of instructional activities were made by the experimenter and teachers cooperatively, subject to the following criteria.

1. The order of the activities as specified is to be rigidly observed, though some activities may be treated lightly or omitted. If they are omitted in AS they are to be omitted also in SA and I, etc.
2. The number of activities completed in any day by any child may vary.
3. The entire instructional period is to be devoted to the given activities. Specifically, new activities will not be introduced into the sequence, nor in general will students be allowed to work in other subject areas during this time.

4. Instruction will stop when all three teachers agree that the activities have been completed to the greatest advantage of the students, but this will not be before four weeks nor later than six weeks after beginning instruction.

Some minor departures from these criteria during the course of the study are noted in the following chapter.

Management decisions peculiar to the individual groups are discussed below.

Group I Management

The activities within I are to be selected in the indicated order as they are needed. However, in no case is a child to be assigned to work on only the addition or subtraction parts of an activity.

Management of Groups AS and SA

The activities of both Treatments AS and SA are also to be assigned in the indicated order. For group AS the addition activities are to continue until two successive P observations (discussed later) indicated group mastery (.75), the addition activities are exhausted, or at the end of three instructional weeks (15 instructional days); the subtraction activities are to begin as soon as possible thereafter. Group SA is to begin and continue with the subtraction activities until two successive P observations indicate group mastery, the subtraction

activities are exhausted, or at the end of three instructional weeks (15 instructional days). The addition activities are to begin as soon as possible thereafter, and continue until instruction is stopped.

Observations

The tests used at the observations 01, P, 02 and 03 are discussed in the order of their administration. Further details of administration and scoring following the descriptions of the instruments. Copies of all the tests used and the estimated reliabilities are found in Appendix C.

The Observation 01

The first part of the test at 01 is an initial measure of student facility with addition and subtraction number facts. The purpose of this part of the test is to provide scores of proficiency with these simple sentences so that any improvement due to incidental drill during the learning of the algorithms can be identified. These scores can also be used as covariates in the analysis of other test data. To facilitate these purposes differences will be maximized by scoring the number correct in three minutes.

The content of this part of 01 is the 30 most difficult addition facts and the 30 most difficult subtraction facts as reported by Suppes, Jerman, and Brian (1968). These items are arranged four items to a line in vertical or horizontal format in a random manner subject to the constraint that no more than two items of any form or

kind occur consecutively. This is true if the test is traversed vertically or horizontally.

The second part of O1 is a measure of student ability to produce sums or differences to two two-digit whole numbers presented in compact vertical form. Scores obtained from this part will provide initial measures of student performance and can also serve as covariates in the analysis of later tests. However, the immediate purpose of this part of the test is to identify subjects who must be dropped from the study because they already possess an adequate addition or subtraction algorithm.

The content of the second part of O1 is thirty addition and subtraction items identified in four subscales. These items were randomly selected and ordered on three forms subject to the following constraints. Every item contains two two-digit numbers and no sum exceeds 99. Every group of six consecutive items consists of two addition items that require regrouping, two subtraction items that require regrouping, one addition item that does not require regrouping, and one subtraction item that does not require regrouping. There are accordingly fifteen addition items, ten of which require regrouping, and fifteen subtraction items, ten of which require regrouping. No more than two items of any kind occur consecutively. Six minutes are allowed for the completion of this part of the test.

Achievement Monitoring Observations

The set P is a set of eight-item forms that are administered

periodically during instruction to provide estimates of group achievement levels. There are nine such forms, each of which contains two randomly selected items suitable for each of the four subscales of the second part of O1 discussed above. All of these items are expressed in vertical form, and one of the addition and one of the subtraction items that require regrouping are expressed in expanded notation. The eight items are randomly ordered on the form subject to the following constraints. The first two items must be one addition and one subtraction item that do not require regrouping. No more than two addition or two subtraction items appear consecutively.

The nine forms of set P are randomly ordered and are assigned to the children in such a way that each form is used with approximately equal frequency at each administration, and that no child will see the same form a second time before he has seen the other forms once. The set of forms is to be administered on the second day of instruction, and again on the second and last days of every week throughout the instructional period. The forms are administered during the first of the period and three minutes are allowed for their completion.

It might be hypothesized that working on these forms twice a week would interact with treatment to affect achievement level and so should be viewed as part of the treatments. While it seems doubtful that 6 minutes per week working on forms that are never seen again, and from which the child receives no direct feedback would be a very powerful influence upon achievement, the strongest evidence against

such an effect is provided by Romberg and Wiles (1972). They report the feasibility of using a very similar procedure for monitoring achievement. In that study a part of a second grade class was randomly chosen to take a set of 5-item test forms every third instructional day during the course of an instructional unit dealing with the addition and subtraction algorithms. A test given at the end of 17 instructional days failed to indicate any differences in scores between the part of the class that had taken the test forms and the part that did not.

The Observation 02

The test at observation 02 is a three part, paper and pencil test. It is to be administered the day following the termination of the instructional treatment. The first two parts are identical to the two sections of 01 described earlier. The third part is a test of the behaviors of tasks A and S. This section is a power test of these behaviors and adequate time is allowed for its completion. This part of 02 contains a four item scale for each behavior measured. Though this part of the test will be referred to as a test of the behaviors associated with the two algorithms, it can also be thought of as a measure of understanding.

The Observation 03

The tests of observation 03 is a repeated administration of the test of 02 with an additional scale in part three. The purpose of this observation

is to measure retention. The added scale will be referred to as a transfer scale. This transfer scale consists of ten items that are extensions of the algorithms to be learned. The observation 03 is to be administered two to three weeks after instruction is stopped.

Administration and Scoring

Observations 01, 02, and 03 are administered by the experimenter in large group settings. The set of administrations of P are administered by the teachers of the children. All tests are scored by the experimenter and checked by personnel not involved with teaching the children. The teachers or children are not informed of any scores except as may be necessary to make management decisions.

Analysis

Corresponding to the way in which the data is collected and the subjects are assigned to treatments, the analysis model is a repeated measures multivariate analysis of variance utilizing randomized blocks. This model is used to examine group mean differences for the variables of 02 and 03. The achievement profiles resulting from the administrations of P are analyzed using a similar model. A detailed explication of the analysis is presented along with the particular questions to which it relates in Chapter VIII.

The following chapter deals with the conduct of the study which is in turn followed by chapters dealing with outcomes.

Chapter VII

CONDUCT OF THE STUDY

Subjects

There were three schools that satisfied the criteria of extensive involvement with DMP. In one of these it was also possible to accomplish the random assignment of subjects to the 3 treatments. This school, Stephen Bull Elementary School in Racine, Wisconsin, is one of several schools in Racine designated by that district as an inner-city school. The school is located in a redeveloped area of the inner city. The parents of the children are employed mostly in factory or services vocations.

The subjects for this experiment then, were 29 boys and 24 girls of the K-2 unit at Stephen Bull Elementary School. The unit contained all children of the school designated administratively as kindergarten, first or second graders. This unit has been using DMP materials since September, 1969, so that most of the children used as subjects had experienced only DMP mathematics materials. All 53 of these students had been judged by their teachers to have completed the topics prerequisite to the experimental topic.

The age span of the subjects was from 87 to 120 months with mean age of 96.6 months and median age of 95.75 months. The range covered by 91 months through 100 months accounts for the middle 75% of the population. The entire distribution of ages is found in Table 1.

Table 1

THE AGE DISTRIBUTION OF SUBJECTS

Age in Months	87	89	90	91	92	93	94	95	96	97	98	99	100	104	105	106	110	112	115	120	Total
Number of Boys	0	4	0	0	4	2	2	3	1	4	0	2	3	1	0	0	1	1	1	0	29
Number of Girls	1	1	1	1	2	3	2	0	4	0	3	3	0	0	1	1	0	0	0	1	24
Total	1	5	1	1	6	5	4	3	5	4	3	5	3	1	1	1	1	1	1	1	53

Total Mean = 96.6

Median = 95.75

range of the middle 75% 91 through 100 months

The subjects were assigned to blocks on the basis of IQ and randomly assigned to treatments from the blocks in the manner specified in the previous chapter. The IQ scores of four children were not available at the time they were assigned to treatments; though IQ scores were later obtained, these four children were placed in an IQ block on the basis of the judgment of the teachers working with the children. The results of this assignment were 17 blocks with 3 subjects, and 1 block with 2 students randomly assigned to treatment groups so that group I contained 17 subjects, and groups AS and SA each contained 18 subjects.

Data from four of the original 53 children are not reported for the following reasons. One child was dropped because of unusually good performance on the pretest O1; his scores indicated he had substantial knowledge of the algorithms to be learned. He was the only subject to exhibit such knowledge. Two subjects were dropped, because in the judgment of the teachers, they had personal adjustment difficulties (not directly related to the study) that left them unable to provide reliable data. A fourth student left school on the first day of the study and so did not provide any data. Neglecting any data from these four left reportable data for 23 girls and 26 boys. The blocks at the beginning of treatment then consisted of 14 blocks with 3 subjects, 3 blocks with 2 subjects, and 1 block with 1 subject; there were 16 subjects in group I, 18 in group SA, and 15 in group AS.

Absences during the course of the study left 12 of the 49 subjects with incomplete data. Complete data are available for 13 subjects in group I, 14 subjects in group SA, and 10 subjects in group AS. However,

since all 12 of the subjects were not gone at the same times, the exact numbers of subjects available for each of the several comparisons are given in the chapter dealing with results along with the data pertaining to the comparison in question. All of the above information concerning the IQ's, blocks, and group assignment is summarized in Table 2.

The Administration of the Experiment

This section is an account of the course of the experiment from the initial assignment of students and teachers to groups, to the administration of the final retention test, O3. This section is followed by a report of details of the testing and exigencies of the study.

The treatment period consisted of twenty-six instructional days during the last half of the spring semester of the 1971-1972 school year. As specified by the design, all treatment groups met at the same hour on the same days throughout the treatment period.

A total of four personnel were involved with the instruction of the groups, three certified elementary school teachers, one of whom was the unit leader, and an instructional aide. These people were the members of the K-2 Unit that had been working with the subjects of the experiment for at least the preceding six months. The certified teachers were randomly assigned responsibility for one of the three treatment groups, while the aide worked among all three groups at the direction of the teachers.

The experimenter was present in the school building Tuesday through Thursday of each week during the course of the experiment to be

Table 2

THE IQ'S OF SUBJECTS BY SEX, BLOCK, AND GROUP ASSIGNMENT

	Block	Group I	Group SA	Group AS
BOYS				
	1	115[d]	121	114
	2	120*	104	88*#
	3	101	98	98*
	4	96	97	97*
	5	96	94*	93
	6	91	90*	91*
	7	85	82	82
	8	75	79	75[d]
	9	75	68	65
	10	--	64	66#[d]
GIRLS				
	11	105	112	131
	12	97	98	97
	13	93*	94	96
	14	93#	93*	88
	15	85*	87	86
	16	82	83	84[d]
	17	81	80*	86*#
	18	77	66	75

*subjects with incomplete data

#IQ scores were obtained during the course of the study

[d]subjects dropped from the study

available to the teachers, and to visit the classrooms at unannounced times during the week to verify that the treatments were being carried out as specified. The teachers were instructed to keep a log indicating the date, the instructional day, the activities being worked on that day, and any comments they might have about the entire program. A form was provided for their use in accomplishing this task.

The assignment of subjects to treatment groups was accomplished on a Monday. Since instruction that day was limited to review activities, this day is not counted as an instructional day. The following day is counted as the first instructional day; on this day, O1 was administered and all groups worked on Activity 1 which was concerned with a review of two-digit numeration. The next day the first periodic observation was made, and the activities of the three groups were differentiated according to the specifications of the treatments.

The materials constructed for treatment I consisted of 14 numbered activities and a set of supplementary activity sheets. The supplementary sheets were to be presented to the students after Activity 9 was completed. The corresponding components of the activities of I that formed treatments AS and SA were labeled 1A, 1S, 2A, 2S, ..., 14A, 14S, and Supp A, and Supp S, where the "A" and the "S" indicate an addition or subtraction component respectively. During the writing of the I activities care was taken to choose activities that could be partitioned without rendering either of the partitions instructionally unsound, but as the final arbiter of acceptability was to be the collective judgment

of the experimenter and the teachers of the children, a great deal of overlap and redundancy of content was deliberately present in the activities that were developed.

During the course of the experiment the instructional activities were selected and paced according to the criteria specified in the preceding chapter with minor exceptions; these exceptions involve activities numbered 8 and 10. While group I worked these activities, group SA worked only 8A and 10A, and group AS worked only 10S. In accordance with the specified criteria the activities numbered 6, 7, 11 and 12 were omitted from all treatments.

Groups AS and SA began with one algorithm and proceeded to the next according to one of the following criterion as specified in the design: exhaustion of materials dealing with the algorithm, or an estimated group mastery level of .75 as measured by the periodic observations for two consecutive observations. The .75 mastery level for addition or subtraction was not reached by any group prior to the fifteenth instructional day, nor were the prepared materials exhausted; therefore, both groups made a transition on the sixteenth instructional day.

This common timing of the transitions of the groups AS and SA from one algorithm to the next, coupled with the express willingness of the teachers, permitted the strengthening of the original design by

including the random reassignment of teachers to groups at the time of this transition in instruction. The basic design allows teacher effect, if it exists differentially, to be confounded with treatment effects. Stated in another way, the administrator of the treatment is part of the treatment. It is argued that the specification of the activities and the role of the teacher as implementor of those specifications allows for likely differences due to teacher behaviors only if the teachers or students are inexperienced with, or antagonistic to this approach to mathematics learning. However, the random reassignment of teachers to groups provides some control for teacher effects that was not specified by the original design.

The reassignments were made subject to the single restraint that no teacher would be reassigned to the group they were with initially. Under this restraint of course, any one random reassignment determined the others as well. The reassignments were as follows. The teacher that began with group AS was reassigned to group I, the teacher that began with group I was reassigned to group SA, and the teacher beginning with group SA was assigned to group AS. The reassignments are summarized in Table 3.

Table 3

THE REASSIGNMENT OF TEACHERS ON THE 16TH INSTRUCTIONAL DAY

Teacher	Original Group	Second Group
1	AS	I
2	SA	AS
3	I	SA

Instruction was terminated on the 26th instructional day and O2 was administered the next day. Test O3 was administered one day short of the two weeks specified by the design. The two weeks specified by the design would have required that the test be administered on the day before the last day of school, it seemed wise to avoid testing at that time.

A brief overview of the conduct of the study, summarizing some of the information contained in this section, is found in Table 4. In particular, Table 4 contains information from the teachers' logs specifying the activities being worked, it indicates the timing of all student tests, and it records the timing of experimenter visits to the classrooms.

The Administration of Tests

The periodic observations were administered by the teachers in the first three minutes of the specified class periods. If a child was absent for one of these administrations, he was tested at the first opportunity upon his return to school. If he missed more than one consecutive administration, only the last missed form was administered; other scores were classified as missing data.

Observations O1, O2, and O3 were all administered in the gymnasium by the experimenter with the aid of the staff involved with these children. As the gymnasium also served as a lunch room for the school the children were seated at the tables in such a way that copying would be obvious.

Table 4

AN OVERVIEW OF THE CONDUCT OF THE STUDY

DAY	DATE	EXPERIMENTAL DAY	OBSERVATIONS	ACTIVITIES OF: GROUP AS	GROUP SA	GROUP I	COMMENTS
M	4/17	0					Assign to groups
Tu	4/18	1	O1	1	1	1	Review numeration
W	4/19	2	P1	2A	2S#	2	
Th	4/20	---					No arithmetic taught
F	4/21	3	P2	2A	2S	2	
M	4/24	4		3A	2S#	2	
Tu	4/25	5	P3	3A*#	3S#	2	
W	4/26	6		3A	3S*#	2	
Th	4/27	7		4A	2S	3*	
F	4/28	8	P4	5A	2S	3	
M	5/1	9		5A	4S	4	
Tu	5/2	10	P5	5A	5S	4*	
W	5/3	11		9A*	5S	5	
Th	5/4	12		9A	9S,SuppS*	5	
F	5/5	13	P6	SuppA	9S,SuppS	8	
M	5/8	14		SuppA	13A,SuppS	8	
Tu	5/9	15	P7	14A	14S,SuppS	9*	
W	5/10	16		2S	2A	9*	Reassignment of teachers
Th	5/11	17		2S*	3A	9	
F	5/12	18	P8	2S	4A	9 Supp	
M	5/15	19		3S	5A	Supp	
Tu	5/16	20	P9	3S	8A*	Supp	
W	5/17	21		4S	9A	Supp*	
Th	5/18	22		5S*	9A,SuppA	Supp	
F	5/19	23	P10	5S	10A,SuppA	10	
M	5/22	24		9S	14A,SuppA	10	
Tu	5/23	25	P11	SuppS	SuppA*	14	
W	5/24	26	P12	SuppS*	SuppA	14	
Th	5/25	--	O2	---	---	--	No instruction relating to algorithms or numeration
W	6/7	--	O3	---	---	--	

* -- the experimenter visited the classroom

\# -- the regular teacher was absent from school

The tests were presented to the children as number problems that they were to do as well as they could. They were encouraged to work those that they could, and to skip those that they could not. The administration of the timed sections included the statement to the children that they would probably not have time to complete all of the items.

The observation O1 was administered shortly after lunch at the beginning of the arithmetic period. A total time of less than thirty minutes was required out of the classroom for the administration of this test. Eight children were absent at this administration: five from group AS and three in group SA; these children were tested by the experimenter during the arithmetic period after they had returned to school. Four of the eight were tested the following day (Day 2), three the following Tuesday (Day 5), and one was dropped from the study.

Observations O2 and O3 were both administered in the same manner. Each administration required two time periods of the same day, one in the morning and the other in the afternoon. Parts one and two along with a few scales of the third part, were administered in the morning period; the total time required for the morning testing period was less than 35 minutes. The remaining scales of the third part were administered in the afternoon period and required less than 20 minutes for the children to complete. Many children required less than 10 minutes.

Exigencies of the Study

Studies that utilize the environment of the operating elementary school are subject to events that cannot be controlled by design considerations. The study being reported is no exception. The only practical alternative to control is to argue that such events happen randomly across groups, and to record them as they occur. Such events do of course place qualifications upon the generalizability of a study, and must be kept in mind when considering the outcomes of any such study. The following paragraphs report such events.

During the first half of the study outbreaks of mumps and measles occurred in the school population causing an unusually high rate of absenteeism during that time. An indication of the extent of this is seen from the number of missing scores for the periodic observations. The mean number of missing scores per group for the first six observations is 1.83; this mean is .89 for the last six observations. The modal number of missing scores per group for the first six observations is 2; the modal number for the last six is 1. The distribution of missed scores by group by periodic observation is shown in Table 5. The distribution supports the judgment of the experimenter that no one group was more severely affected than another. However, such absences may well have worked to hold down total achievement scores.

A potentially more serious problem concerns teacher absences. Absences of the certified teacher with assigned responsibility for a treatment group are indicated in Table 4. One such absence occurred on

Table 5

MISSING DATA FOR EACH GROUP AT EACH PERIODIC OBSERVATION

Day	2	3	5	8	10	13	15	18	20	23	25	26
Observation	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Group												
AS	2	4	2	1	2	1	3	1	1	1	0	0
SA	2	2	1	2	2	1	0	1	1	1	1	1
I	1	2	1	3	2	2	0	1	1	1	1	1

the fifth day for Treatment AS, and absences occurred on instructional days 2, 4, 5 and 6 for Treatment SA. The absences associated with Treatment SA were due to illness. Though experienced substitute teachers were provided, the instructional aide assumed primary responsibility for administering the treatment on all but two of these occasions. The experimenter met with the substitutes to detail exactly what was involved with carrying out the activities selected for those days. The judgment of all involved was that this illness and the associated adjustments necessitated by it may have resulted in the net loss of one day's instruction with group SA. While it is impossible to assess precisely the effects of this set of events, the fact that they occurred early in the treatment period make it reasonable to assume that achievement differences attributable to them become progressively less likely throughout the treatment period. Though this matter will be mentioned again later, it is generally assumed that no differences are

attributable to this set of events past the tenth day.

The purpose of this chapter has been to identify the subjects of the study, relate the administration of the study to the design, and to describe those unpredictable events that could affect interpretations of the outcomes. The final chapters of this report deal with the specifications of outcomes, tests of hypotheses, and a discussion of the questions of this study in terms of that information.

Chapter VIII

DATA DESCRIPTION AND ANALYSIS

The preceding chapter described the collection of data relating to the questions of the study. This chapter reports the hypotheses, the data collected to test the hypotheses, the statistical model employed to analyze the data, and the results of the tests of these hypotheses.

This chapter begins with a resume of the questions of the study and a statement of the hypotheses. All of the variables and the labels that are used to identify the data are listed in the next section. Following that the statistical model used to estimate the group means and to test the hypotheses of the study is described; the concluding sections report the data of the observations and the tests of the hypotheses as well as a summary of the results of the study.

The Hypotheses

The questions of this study concern the effects of three instructional sequences on children's achievement. The sequences were designed to teach the two-digit addition and subtraction algorithms. Though a complete listing of the questions is given in

Chapter V, the primary questions concern the children's performance immediately following instruction and their performance some time after instruction. Also of interest are the rate at which the children acquire performance with the algorithms and the effects of the instructional treatments on the children's speed with addition and subtraction facts.

Data were gathered to test the following hypotheses.

H1: There are no differences in group performance during instruction.

H2: There are no differences in group performance on timed tests of achievement.

H3: There are no differences in group performance on power test measures of total performance.

H4: There are no differences in group performance on power test measures of differential performance for addition and subtraction.

The data which were used to test these hypotheses came from standardized test scores obtained from school records, the initial observation O1, twelve periodic observations made during instruction, the observation O2 taken immediately following the instructional period, and the observation O3 taken two weeks after instruction was completed. Both achievement measures in which speed is a relevant factor and power test measures of achievement, in which it is not, were obtained from observations O2 and O3. Measures of the rate of forgetting were constructed from the O2 and O3 data; these measures

are referred to as 03 - 02 data. The above sets of data relate to the hypotheses as follows.

Hypothesis H1 was broken down into several sub-hypotheses that were tested by the periodic observation data. These sub-hypotheses are presented with the report of the periodic observation data. Each of the hypotheses H2, H3, and H4 were tested by data from 02, 03 and 03 - 02. The tests of hypotheses H2, H3, and H4 provided by 02 data describe performance following instruction, the tests provided by 03 data describe retention, and the tests provided by 03 - 02 describe the rate of forgetting. The actual null hypotheses tested by the data from each observation are stated with the report of the data from that observation.

Before preceding with the reporting of the data and the tests of the hypotheses, the next section summarizes the observations and variables which provide the data.

The Specification of Variables

The purpose of this section is to summarize the observations and to describe the variables measured at those observations. In addition, this section specifies the labels used to identify the data. Finally, this information is summarized in Table 6.

Observations

Tests were administered at 15 different times as summarized in Table 4 of Chapter VII; they were observations 01, 02, and 03, and the 12 periodic achievement monitoring observations. These 15

Table 6

The Observations and Variables

Label	Description	Range for P1-P12	Range for 01, 02, 03
Observations			
C1	Pre-treatment Performance		
02	Post treatment Performance		
03	Retention Performance		
P1 through P12	Second and Last Day of Each Week Observations		
T1	(P1 + P2 + P3 + P4)/4		
T2	(P5 + P6 + P7)/3		
T3	(P8 + P9)/2		
T4	(P10 + P11 + P12)/3		
03 - 02	Variable values at 02 subtracted from values at 03		
Status Variables			
IQ	Otis-Lennon Mental Ability Test Deviation I.Q.		
MET1	Computation Score - Metropolitan Achievement Test		
MET2	Concepts Score		
MET3	Problem Solving Score		
Measured Achievement Variables			
+F	Addition Facts Speed Test		0,30
-F	Subtraction Facts Speed Test		0,30
ANC	2-digit addition, no carrying	0,2	0,5
SNB	2-digit subtraction, no borrowing	0,2	0,5
C	2-digit addition, carrying	0,2	0,10
B	2-digit subtraction, borrowing	0,2	0,10
Constructed Achievement Variables			
SUM	* (ANC + C/2 + SNB + B/2) /4		0,5
	# (ANC + C + SNB + B) /4	0,4	
OP	* (ANC + C/2 - SNB - B/2) /2		-5,5
	# (ANC + C - SNB - B) /2	-4,4	
RGP	* (ANC - C/2 + SNB - B/2) /2		-5,5
	# (ANC - C + SNB - B) /2	-4,4	
INT	* (ANC - C/2 - SNB + B/2)		-5,5
	# (ANC - C - SNB + B) /2	-4,4	

* relates to measures at 01, 02, and 03

\# relates to measures at P1 through P12

Table 6 (continued)

Measured Power Test Variables

Label	Description	Range
AR	Addition Regrouping	0,4
SR	Subtraction Regrouping	0,4
ANCE	Addition - no carrying - expanded notation	0,4
SNBE	Subtraction - no borrowing - expanded notation	0,4
CE	Addition - carrying - expanded notation	0,4
BE	Subtraction - borrowing - expanded notation	0,4
C+E	Writes compact as expanded - addition	0,4
C-E	Writes compact as expanded - subtraction	0,4
ANCC	Addition - no carrying - compact notation	0,4
SNBC	Subtraction - no borrowing - compact notation	0,4
CC	Addition - carrying - compact notation	0,4
BC	Subtraction - borrowing - compact notation	0,4
A(10)	Addition - tens only - compact notation	0,4
S(10)	Subtraction - tens only - compact notation	0,4

Constructed Power Test Variables

Total Power Test Variables		
SUMB1	AR + SR	0,8
SUMB2	ANCE + SNBE	0,8
SUMB3	CE + BE	0,8
SUMB4	C+E + C-E	0,8
SUMB5	ANCC + SNBC	0,8
SUMB6	CC + BC	0,8
SUMB7	A(10) + S(10)	0,8
Difference Power Test Variables		
DIFB1	AR - SR	-4,4
DIFB2	ANCE - SNBE	-4,4
DIFB3	CE - BE	-4,4
DIFB4	C+E - C-E	-4,4
DIFB5	ANCC - SNBC	-4,4
DIFB6	CC - BC	-4,4
DIFB7	A(10) - S(10)	-4,4

administrations of tests yielded 15 sets of variable values. From these 15 sets, 5 other sets of values were constructed.

One of the constructed sets is the difference scores between the 02 and 03 variable values, 03 - 02. The other constructed sets of values were determined from the 12 periodic observations; this determination is specified with the report of the achievement monitoring data. It is sufficient at this point to identify them as values for the time periods T1, T2, T3, and T4. These five sets of derived values, 03 - 02, T1, T2, T3, and T4, along with the original 15 sets of data are referred to as "the observations." There are then a total of 20 observations referred to in the report and analysis of data.

The tests administered at the various observations are described in general terms in Chapter VI. Some, but not all, of the variables measured by those tests were specified there. The following paragraphs identify all of the variables, put them in appropriate categories, give them a label, and provide other pertinent information about those variables.

Basically, the variables are considered in three categories: status variables measured by standardized test scales, achievement variables measured under timed conditions, and power test variables measured under conditions where time is not an important factor.

Both achievement and power test variables were measured at the observations 02 and 03, while only achievement variables were measured at the observation 01 and at each of the periodic observations.

while the achievement variables of 01, 02, and 03 were measured by exactly the same items, the achievement variables of the periodic observations were measured by other and fewer items. However, these items measured some of the same variables as did the items used in 01, 02, and 03. The following listing is broken down into sections that separate the variables of the periodic observations from those of the other observations.

Both the achievement variables and the power test variables referred to above were measured by tests developed for use in this study. These two categories of variables contain those that were measured directly by a specified set of items, and those for which a score was constructed by a linear combination of the values that were combined for directly measured variables. The directly measured variables of each category are listed first.

The specification of the directly measured variables includes a label, a descriptive phrase, the number of items used to measure the variable, a description of the items that were used, and an example of an item. The specification of the constructed variables consists of a label, the possible range of the scores for the variable, and a description of how the value for the variable was obtained.

When a variable is referred to in the reporting of the data, the variable label will always be used, however, since most of the variables were measured at several observations, a code consisting of the observation label and the variable label is used in any

situation where confusion might otherwise result. For example, 02ANC refers to the variable ANC measured at observation 02.

The order of the following listing is: the status variables, the variables of the periodic observations, the other achievement variables, and the power test variables.

Status Variables

There are four status variables whose values were determined by school administered standardized tests. The first of these is a measure of mental ability (IQ), and the other three are the three mathematics scales of the Metropolitan Achievement Test (1970) that was administered by the school district the week prior to the beginning of the experiment. The values of these four variables were determined as follows.

IQ — Deviation IQ This is the Deviation IQ obtained from the Otis-Lesson Mental Ability Test (1968).

MET1 — Computation This is the raw score of the Computation scale of the Metropolitan Achievement Test.

MET2 — Concepts This is the raw score of the Concepts scale of the Metropolitan Achievement Test.

MET3 — Problem Solving This is the raw score of the Problem Solving scale of the Metropolitan Achievement Test.

Periodic Observation Achievement Variables

Directly measured variables. Each of the forms used to measure periodic achievement contained eight items as specified below. An item was scored correct or incorrect, and the variable value for each

of the directly measured variables is simply the number of correct items that relate to that variable. There are four directly measured periodic observation achievement variables.

ANC Addition No Carrying (2 items) Each item is two two-digit whole numbers written in vertical form and whose sum is found without regrouping.
EXAMPLE:

$$\begin{array}{r} 71 \\ +\ 26 \\ \hline \end{array}$$

C Addition With Carrying (2 items) Each item is two two-digit whole numbers written in vertical form and whose sum is found by regrouping ones. One of these items is expressed in compact notation and the other in expanded notation.
EXAMPLES:

$$\begin{array}{r} 79 \\ +\ 18 \\ \hline \end{array} \qquad \begin{array}{r} 1(10) + 4 \\ +\ 1(10) + 8 \\ \hline \end{array}$$

SNB Subtraction No Borrowing (2 items) Each item is two two-digit whole numbers written in vertical form and whose difference is found without regrouping.
EXAMPLE:

$$\begin{array}{r} 56 \\ -\ 32 \\ \hline \end{array}$$

B Subtraction With Borrowing (2 items) Each item is two two-digit whole numbers written in vertical form and whose difference is found by regrouping tens. One of these items is expressed in compact notation and the other in expanded notation.
EXAMPLES:

$$\begin{array}{r} 96 \\ -\ 38 \\ \hline \end{array} \qquad \begin{array}{r} 8(10) + 5 \\ -\ [7(10) + 8] \\ \hline \end{array}$$

Constructed measures periodic achievement variables. Five variables have values constructed from the values for the preceding four periodic observation achievement variables. The values for the first of these were determined only for the 12 periodic measures, and the values of the last four were determined only for four consolidated time periods. The five variables are as follows:

Total — Total Score (Range 0 to 8) An individual's score for this variable was formed by adding his scores for the variables ANC, C, SNB, and B. This variable was computed only for the 12 periodic observations.

SUM — Algorithm Achievement Score (Range 0 to 2) An individual's score for this variable was formed by adding his scores for the variables ANC, C, SNB, and B, and then dividing by four.

OP — Operation Difference (Range -2 to 2) An individual's score for this variable was formed by adding his scores for ANC and C and then subtracting his scores for SNB and B. The result was then divided by two.

RGP — Regrouping Difference (Range -2 to 2) An individual's score for this variable was formed by adding his scores for ANC and SNB and then subtracting his scores for C and B. The result was then divided by two.

INT — Interaction (Range -2 to 2) An individual's score for this variable was formed by adding his scores for ANC and B and then subtracting his scores for C and SNB. The result was then divided by two.

Achievement Variables of Observations 01, 02, and 03

Directly measured achievement variables. There are six achievement variables which were directly measured at observations 01, 02, and 03. An individual's score for each of these variables was the number of correct responses for the items that related to that variable. These variables are as follows:

+F — Addition Facts (30 items). Each item is one of the 30 most difficult addition facts as identified in Chapter VI. Fifteen of these items are expressed in vertical form and fifteen in horizontal form.
EXAMPLES:

$$\begin{array}{r} 6 \\ +\ 3 \\ \hline \end{array} \qquad 8 + 5 =$$

-F — Subtraction Facts (30 items). Each item is one of the 30 most difficult subtraction facts as identified in Chapter VI. Fifteen of these items are expressed in vertical form and fifteen in horizontal form.
EXAMPLES:

$$\begin{array}{r} 12 \\ -\ 5 \\ \hline \end{array} \qquad 7 - 5 =$$

ANC Addition No Carrying (5 items). Each item is two two-digit numbers written in vertical form and whose sum is found without regrouping.
EXAMPLE:

```
  22
+ 56
```

C Addition With Carrying (10 items). Each item is two two-digit whole numbers written in vertical form and whose sum is found by regrouping ones.
EXAMPLE:

```
  18
+ 33
```

SNB Subtraction No Borrowing (5 items). Each item is two two-digit whole numbers written in vertical form and whose difference is found without regrouping.
EXAMPLE:

```
  95
- 84
```

B Subtraction With Borrowing (10 items). Each item is two two-digit whole numbers written in vertical form and whose difference is found by regrouping tens.
EXAMPLE:

```
  54
- 18
```

Constructed measures of 01, 02, and 03 achievement variables.

Four variables have values constructed from the values for the 01, 02, and 03 directly measured achievement variables ANC, C, SNB, and B. Values for these variables were determined as follows:

SUM Algorithm Achievement Score (Range 0 to 5) An individual's score for this variable was formed by adding his scores ANC and SNB to half of his scores for B and C, and then dividing the sum by 4. (ANC + C/2 + SNB + B/2) / 4

OP Operation Difference (Range -5 to 5) An individual's score for this variable was formed by adding his score for ANC to half of his score for C and then subtracting his score for SNB and half of his score for B. The result was then divided by two.
(ANC + C/2 - SNB - B/2) /2

RGP Regrouping Difference (Range -5 to 5) An individual's score for this variable was formed by adding his scores for ANC and SNB and then subtracting half of his scores for C and B. The result was then divided by two. (ANC -C/2 + SNB - B/2) /2

INT Interaction (Range -5 to 5) An individual's score for this variable was formed by adding his score for ANC and half of his score for B and then subtracting his score for SNB and half of his score for C. The result was then divided by two. (ANC - C/2 = SNB + B/2) /2

Power Test Variables of Observations 02 and 03

Directly measured power test variables. There are 14 directly measured power test variables. These are identified in pairs as they relate to similar aspects of the addition and subtraction algorithms. These power test variables, which do not directly measure the behaviors given in the behavioral hierarchy, can be associated with those behaviors; that association is indicated in Appendix A. The following listing presents the variables in paired order. An individual's score is the number of items correctly completed for that variable.

AR Addition Regrouping (4 items) Each item is an open grouping sentence that is solved by completing the regrouping associated with carrying.
EXAMPLE:

$$3(10) + 14 = \_\_\_\_\_(10) + 4$$

SR Subtraction Regrouping (4 items) Each item is an open grouping sentence that is solved by completing the regrouping associated with borrowing.
EXAMPLE:

$$3(10) + 4 = 2(10) + \_\_\_\_\_$$

ANCE Addition No Carrying: Expanded Notation (4 items) Each item is two two-digit whole number written in vertical form and expanded notation whose sum is found without regrouping.
EXAMPLE:

$$\begin{array}{r} 2(10) + 4 \\ +\ 3(10) + 5 \\ \hline \end{array}$$

SNBE Subtraction No Borrowing: Expanded Notation (4 items) Each item is two two-digit whole numbers written in vertical form and expanded notation whose difference is found without regrouping.
EXAMPLE:

$$\begin{array}{r} 4(10) + 8 \\ -\ 2(10) + 5 \\ \hline \end{array}$$

CE Addition with Carrying: Expanded Notation (4 items) Each item is two two-digit whole numbers written in vertical form and expanded notation whose sum is found by regrouping ones.
EXAMPLE:

$$\begin{array}{r} 2(10) + 8 \\ +\ 3(10) + 9 \\ \hline \end{array}$$

BE Subtraction With Borrowing: Expanded Notation (4 items) Each item is two two-digit whole numbers written in vertical form and expanded notation whose difference is found by regrouping tens.
EXAMPLE:

$$\begin{array}{r} 4(10) + 3 \\ -\ 2(10) + 5 \\ \hline \end{array}$$

C+E Rewrites Compact Notation Addition Example (2 items; Range 0 to 4) Each item consists of a compact notation addition example that the child is to write in expanded notation. One point is scored for each number correctly rewritten.
EXAMPLE:

$$\begin{array}{r} 34 \\ +\ 27 \\ \hline \end{array} \quad \text{is the same as} \quad \begin{array}{r} \_\_(10) + \_\_ \\ +\ \_\_(10) + \_\_ \\ \hline \end{array}$$

C-E Rewrites Compact Notation Subtraction Example (2 items; Range 0 to 4) Each item consists of a compact notation subtraction example that the child is to write in expanded notation. One point is scored for each number correctly rewritten.
EXAMPLE:

$$\begin{array}{r} 45 \\ -\ 18 \\ \hline \end{array} \quad \text{is the same as} \quad -\left[\begin{array}{r} \_\_(10) + \_\_ \\ \_\_(10) + \_\_ \end{array}\right]$$

ANCC Addition No Carrying: Compact Notation (4 items) Each item is two two-digit whole numbers written in vertical and compact notation whose sum is found without regrouping.
EXAMPLE:

$$\begin{array}{r} 34 \\ +\ 23 \\ \hline \end{array}$$

SNBC Subtraction No Borrowing: Compact Notation (4 items) Each item is two two-digit whole numbers written in vertical form and compact notation whose difference is found without regrouping.
EXAMPLE:

```
  49
- 32
```

CC Addition With Carrying: Compact Notation (4 items) Each item is two two-digit whole numbers written in vertical form and compact notation whose sum is found by regrouping ones.
EXAMPLE:

```
  53
- 29
```

A(10) Addition: Tens Only (4 items) Each item is two two-digit whole numbers written in vertical form and compact notation whose sum is to be found. Both of the two digit numbers have zero in the ones place.
EXAMPLE:

```
  30
+ 40
```

S(10) Subtraction: Tens Only (4 items) Each item is two two-digit whole numbers written in vertical form and compact notation whose difference is to be found. Both of the two-digit numbers have zero in the ones place.
EXAMPLE:

```
  80
- 50
```

Constructed measure power test variables. The scores of the 14 paired variables do not enter directly into the analysis of the group differences, and their means are not reported. However, a set of seven variables composed of the sums of the seven paired variable values is used to measure total performance for the power test variables, while the differences of the seven paired variable values provide measures of an individual's performance for addition over subtraction. The first seven variables are referred to as the total power test variables,

and the second seven as the difference power test variables. They are listed in that order.

Total Power Test Variables

SUMB1 Regrouping (Range 0 to 8) The score is formed by adding the scores for AR and SR.

SUMB2 Expanded Notation: No Regrouping (Range 0 to 8) The score is formed by adding the scores for ANCE and SNBE.

SUMB3 Expanded Notation: Regrouping (Range 0 to 8) The score is formed by adding the scores for CE and BE.

SUMB4 Compact to Expanded Notation (Range 0 to 8) The score is formed by adding the scores for C+E and C-E.

SUMB5 Compact Notation: No Regrouping (Range 0 to 8) The score is formed by adding the scores for ANCC and SNBC.

SUMB6 Compact Notation: Regrouping (Range 0 to 8) The score is formed by adding the scores for CC and BC.

SUMB7 Tens Only (Range 0 to 8) The score is formed by adding the scores for A(10) and S(10).

Difference Power Test Variables

DIFB1 Difference Regrouping (Range -4 to 4) The score is formed by subtracting the score for SR from that for AR.

DIFB2 Difference Expanded Notation: No Regrouping (Range -4 to 4) The score is formed by subtracting the score for SNBE from that for ANCE.

DIFB3 Difference Expanded Notation: Regrouping (Range -4 to 4) The score is formed by subtracting the score for BE from that for CE.

DIFB4 Difference Compact to Expanded Notation (Range -4 to 4) The score is formed by subtracting the score for C-E from that for C+E.

DIFB5 Difference Compact Notation: No Regrouping (Range -4 to 4) The score is formed by subtracting the score for SNBC from that for ANCC.

DIFB6 Difference Compact Notation: Regrouping (Range -4 to 4) The score is formed by subtracting the score for BC from that for CC.

DIFB7 Difference Tens Only (Range -4 to 4) The score is formed by subtracting the score for S(10) from that for A(10).

A summary of the information provided by this section is given in Table 6.

The status variables were measured by commercially available standardized tests; the scales for these variables were therefore assumed to be face valid, and no estimates of the reliabilities were made. The other variables used, the non-status variables, define sets of items which were sampled to obtain the test scales used; these tests therefore are content valid. Because the number of items used to measure each non-status variable was low, each scale contained from four to ten items, no reliability estimates were generated for these scales either. The reliabilities, however, could be expected to be low. If so, this would result in a loss of power to detect true differences and hence conservative tests of group differences.

The Data Analysis Model Employed

The raw scores of each subject for the variables just described were used to obtain estimated group means for the variables. While some of the means reported are the observed group means for these variables, all of the analyses to test for group differences used the estimated group means. This section presents the model used to generate the estimated group means and outlines the plan employed to analyze the data.

The basic model used in the analysis is an incomplete randomized block design of 18 blocks and three treatments. The 18 blocks are the IQ blocks identified in Chapters VI and VII, and the treatments are AS, SA, and I. This basic design has been explicated by Scheffé (1959). Each score, y_{ij}, is assumed to be the sum of a grand mean, μ, a constant block effect b_i, a constant treatment effect a_j, and an error, e_{ij}. The set of errors are assumed to be independently and identically, normally distributed random variables with mean 0 and variance σ_e^2. This model is represented by the following equation.

$$y_{ij} = \mu + b_i + a_j + e_{ij}$$

The constants b_i and a_j are estimated by least squares analysis so that tests involving the a_j's are unbiased by any block effects. The analysis of variance F tests of group differences in this model assume the blocks factor is a random factor and the interaction of blocks and treatments is zero. From these assumptions, the usual F test results have been shown to be conservative (Scheffé, 1959).

Violation of the assumption that the errors are normally distributed with equal variances has been shown to be relatively unimportant as the F tests used are generally insensitive to such violations. This assumption along with the assumption that the error means are identically zero are best dealt with by good experimental control. Steps taken in this study to insure good experimental control were: the random assignment of subjects from blocks to treatments, the treating and testing of all subjects at the same time, and the employment of large group testing. Absences are considered to be random across groups.

The assumption concerning independence of the errors has not been fulfilled; in particular, the subjects of a group were not treated in isolation from the rest of the group. The use of this model in this study then assumes that dependencies among responses are varying in their effects and are generally small so that the overall effect of such dependencies is negligible. Statistical estimates of the dependencies or the empirical effects of such dependencies are not presently available.

The estimated group means produced by the randomized block model are used to describe the performance of the groups at the various observations. However, many of the tests for significant differences are made by adding covariates to the model in the usual manner described by Cochran (1957).

Additional assumptions imposed upon the model by the inclusion of covariates is that the regression be linear and that their values be independent of treatment. The random assignment of subjects to treatments insures the second condition. Linear regression will be assumed for two reasons: (1) it is the usual finding in educational research and (2) no covariate will be used in testing hypotheses if its inclusion does not account for a significant amount of variance in the dependent variables. The procedure for selecting among the potential covariates MET1, MET2, MET3, and O1 variables is given in the section which describes their selection.

Significance Level

The significance level for type 1 errors in educational research has traditionally been .05 or .01. However, as Hays (1963) has

demonstrated, this choice, along with the selection of a statistical model and sample size, automatically affects the type 2 error. As the type 1 error decreases, the type 2 error increases. While it is often true that inferring differences when there are none is of greater portent than failing to infer differences when they do exist, there seem to be no reasons for preferring a type 2 error to a type 1 error in this study. Therefore, some estimates of the power of the tests were made and the following decision regarding the significance level was made. The actual p values obtained for each F ratio are reported. Those F ratios with p values less than .10 are regarded as significant. When the p value for the F test is between .10 and .15, the differences are said to approach significance.

The Status Variables and Observation 01 Data

This section presents the group means for the status variables and the variables of observation 01. While these means provide a description of initial group performances, covariates for use in the analysis of the data from 02 and 03 were also selected from these variables.

Neglecting the data obtained from subjects who were dropped from the study for reasons specified in Chapter VII, the observed group means and the estimated group means for the three treatment groups are reported in Tables 7 and 8. The means for all of these variables,

Table 7

Observed Group Means for 01 Variables and Status Variables

Group (N)	Variables									
	Status Variables				Achievement Variables					
	IQ	MET1	MET2	MET3	+F	-F	ANC	C	SNB	B
AS(15)	92.75	16.3	22.7	18.9	6.20	5.60	1.93	0	.73	.07
SA(18)	89.75	15.8	22.8	18.6	4.39	4.39	1.72	.17	.83	0
I(16)	89.07	15.7	21.9	17.2	4.75	3.94	1.50	.19	1.25	0

Table 8

Estimated Group Means and Standard Errors of 01 Variables and Status Variables

Group (N)	Variables									
	Status Variables				Achievement Variables					
	IQ	MET1	MET2	MET3	+F	-F	ANC	C	SNB	B
AS(15)	89.50	16.0	22.5	18.7	6.13	5.51	1.78	-.03	.71	.07
	1.65	1.30	1.59	1.07	.70	.78	.42	.15	.43	.04
SA(18)	89.69	15.8	22.8	18.6	4.39	4.39	1.72	.17	.83	.00
	1.32	1.13	1.38	.93	.61	.68	.36	.13	.37	.04
I(16)	89.47	16.0	22.2	17.6	4.80	4.01	1.41	.17	1.18	.00
	1.61	1.24	1.52	1.03	.67	.75	.40	.14	.41	.04

NOTE: Entries are of the form (Estimated mean)
(Estimated standard error)

with the exception of those for the variables C and B are greater than zero. Though a few items were done correctly by some Ss, the group means for these two variables are practically zero.

The presence of tally marks on students' paper and the small number of completed items indicate that the nonzero scores for variables B and C are the result of counting by ones to find the sum or difference. Three students correctly completed items for variable C; one student completed one item correctly, another two and a third three. The last two students had a total of three and four correct items, respectively, for all of the variables ANC, C, SNB, and B. The only student to correctly complete an item of the scale for variable B, correctly completed only one other item during the six-minute time interval. In all of these cases the judgment was made that no student used an algorithm, apart from simple counting, to solve the items for the variables B and C.

While there are no experimental hypotheses relating to group performance at observation 01, this data was analyzed for differences at the beginning of instruction that might have occurred in spite of the random assignment of subjects to treatments. The hypothesis that was tested is:

H_o: There are no differences among the estimated group means at 01.

First, this hypothesis was tested using a multivariate analysis of variance on the estimated group means for the 01 variables +F, -F,

ANC, C, SNB, and B; it was further tested by univariate analysis of variance on each of these variables as well as on the variable O1SUM. The F ratios and associated p levels for these ratios are found in Table 9. Since none of the resulting F ratios even approach significance, the hypothesis is not rejected.

Table 9

MANOVA and ANOVAs for O1 Variables

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	1.23	.30	12/48 N/D
+F	12.39	1.88	.18	2/29 hyp/error
-F	8.96	1.08	.36	2/29
ANC	.60	.25	n.s.	2/29
C	.20	.66	n.s.	2/29
SNB	.90	.36	n.s.	2/29
B	.02	1.06	.37	2/29
SUM	.02	.03	n.s.	2/29

The Periodic Achievement Measures

The periodic achievement measures are used to describe the performance of each group during the course of instruction. There were twelve administrations of the eight-item forms as indicated in Chapter VII (see Table 4). Data from these twelve administrations were consolidated into data for four time periods as described later. Two of these time periods, however, cover the periodic observations P1 through P7 and two cover the observations P8 through P12. As was indicated in Chapter VII, the change of content for groups SA and SA and the concomitant change of teachers was effected the day following the observation P7. This section first reports the observed means for the twelve observations P1 through P12, then the estimated group means for the four time periods, and finally the tests relating to hypothesis H1. A summary concludes this section.

The observed group means and standard deviations for the four measured variables of the observations P1 through P12 are reported in Table 10. Figures 2 through 5 provide a graphical display of these means. Since there were two items for each of the measured variables ANC, SNB, C, and B, the possible range of the means for these variables is zero to two. The standard deviations of the means for the variables ANC, SNB, C, and B are quite large in comparison to the means; this is due in part to the small possible range of an individual's score.

Though some shifts in relative position occur, the group means for the variables ANC and SNB show a gradual increase over time that

Table 10

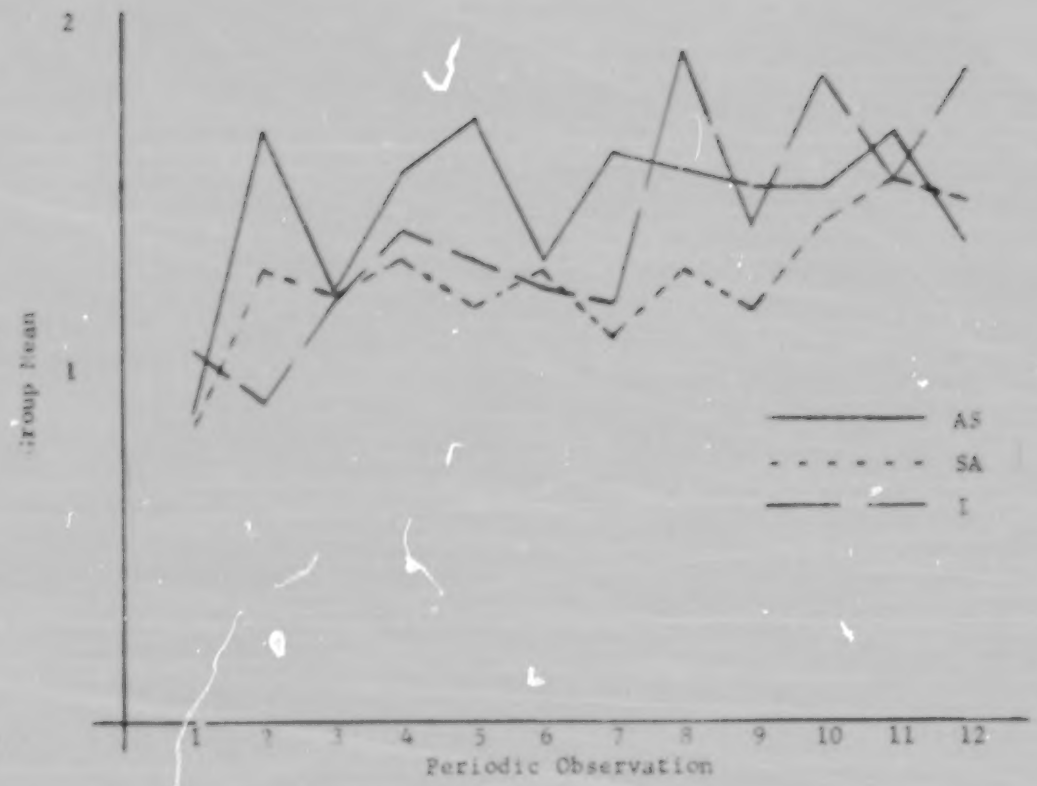
Observed Group Means and Standard Deviations of Variables at the Twelve Periodic Observations

Administration	P1			P2			P3			P4			P5			P6		
Group	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I
Variable																		
ANC	.90	.86	1.07	1.70	1.31	.92	1.25	1.24	1.21	1.58	1.33	1.42	1.73	1.20	1.38	1.33	1.31	1.31
	.94	.87	.80	.67	.79	.95	.75	.90	.89	.79	.82	.79	.65	.77	.77	.89	.87	.85
C	.36	.07	.20	.10	.19	.15	.17	.24	.29	.92	.07	.17	.91	.27	.46	1.33	.25	.23
	.50	.27	.56	.32	.54	.38	.39	.56	.61	.90	.26	.39	.94	.46	.66	.78	.58	.60
SNB	.91	.50	.93	1.10	.94	1.00	1.00	.94	.86	1.25	.93	.92	1.18	1.20	1.08	.75	1.31	.92
	.94	.65	.96	.88	.77	1.00	1.04	.90	.95	.97	1.03	.90	.98	.86	.86	.87	.87	.95
B	0	0	0	0	0	0	0	0	0	0	0	0	0	.13	.15	0	.13	.08
	0	0	0	0	0	0	0	0	0	0	0	0	0	.35	.38	0	.50	.28
No. of the Mean	11	14	15	10	16	13	12	17	14	12	15	12	11	15	13	12	16	13

Table 10 (Continued)

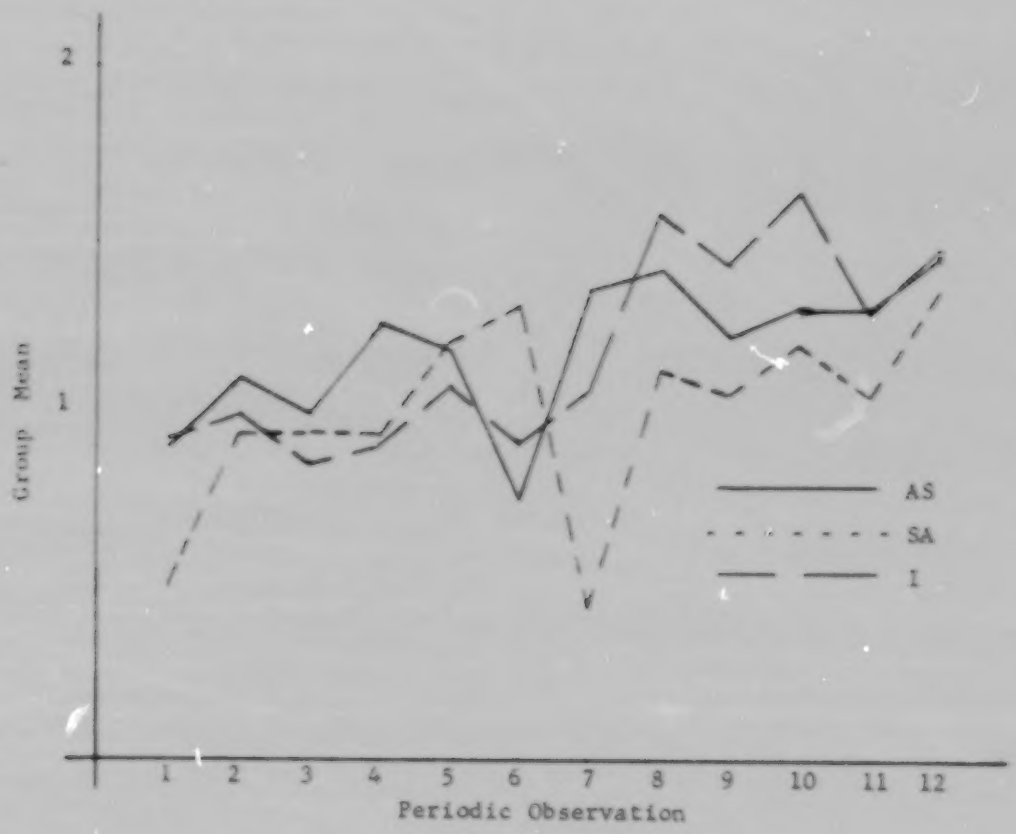
Administration	P7			P8			P9			P10			P11			P12		
Group	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I	AS	SA	I
Variable																		
ANC	1.64	1.12	1.20	1.58	1.31	1.93	1.54	1.19	1.43	1.54	1.44	1.86	1.69	1.56	1.43	1.38	1.50	1.87
	.67	.93	.77	.79	.79	.27	78	.98	.85	.78	.73	.36	.75	.73	.76	.97	.73	.52
C	1.00	.18	.33	1.33	.69	1.00	.77	.81	1.21	1.08	1.13	1.21	1.08	1.00	.93	1.15	1.13	1.13
	1.00	.53	.72	.78	.87	.78	.93	.83	.80	.86	.72	.89	.86	.97	.92	.90	.89	.92
SNB	1.36	.94	1.07	1.42	1.13	1.57	1.23	1.06	1.43	1.31	1.19	1.64	1.31	1.06	1.29	1.46	1.31	1.47
	.92	.90	.88	.67	.81	.76	.83	.93	.85	.95	.91	.63	.85	.77	.83	.88	.79	.83
B	0	.47	.07	.08	.25	.29	0	.50	.50	.54	.50	.93	.77	.44	.57	.77	.50	1.00
	0	.72	.26	.29	.58	.61	0	.73	.52	.78	.82	.92	.93	.73	.76	1.01	.82	1.00
No of the Mean	11	17	15	12	16	14	13	16	14	13	16	14	13	16	14	13	16	15

NOTE: The Entries are of the form Mean
Standard deviation



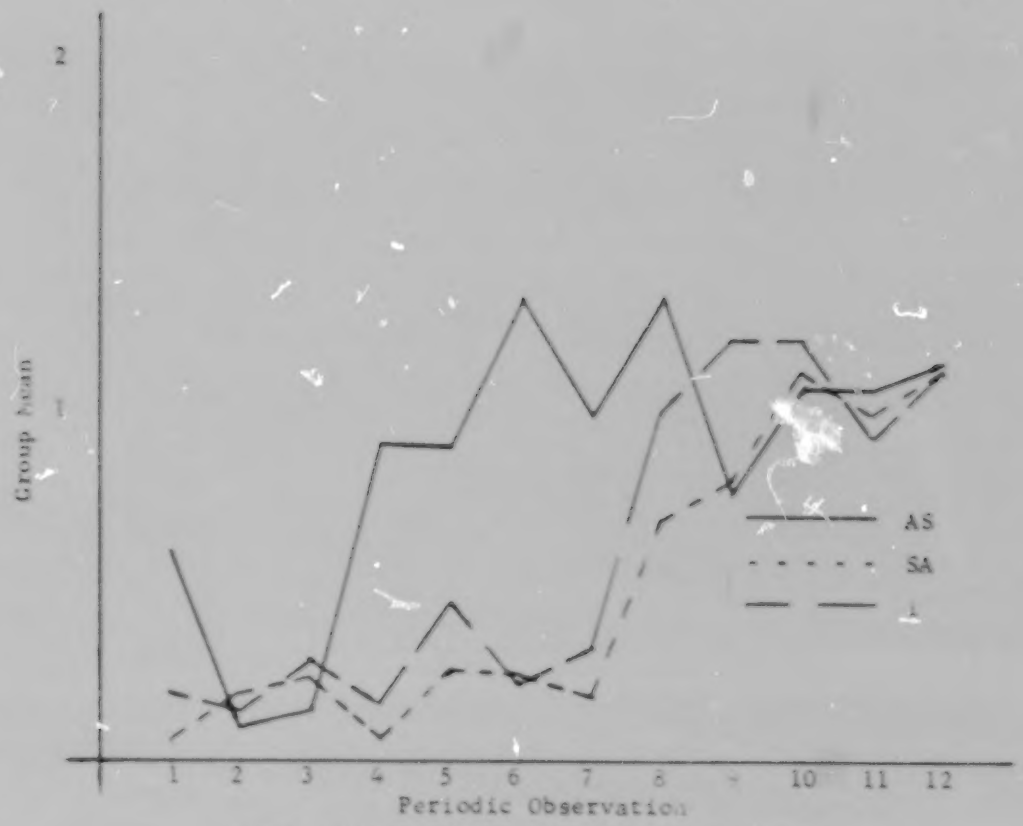
Graphs of Group Means for P1ANC Through P12ANC

Figure 2



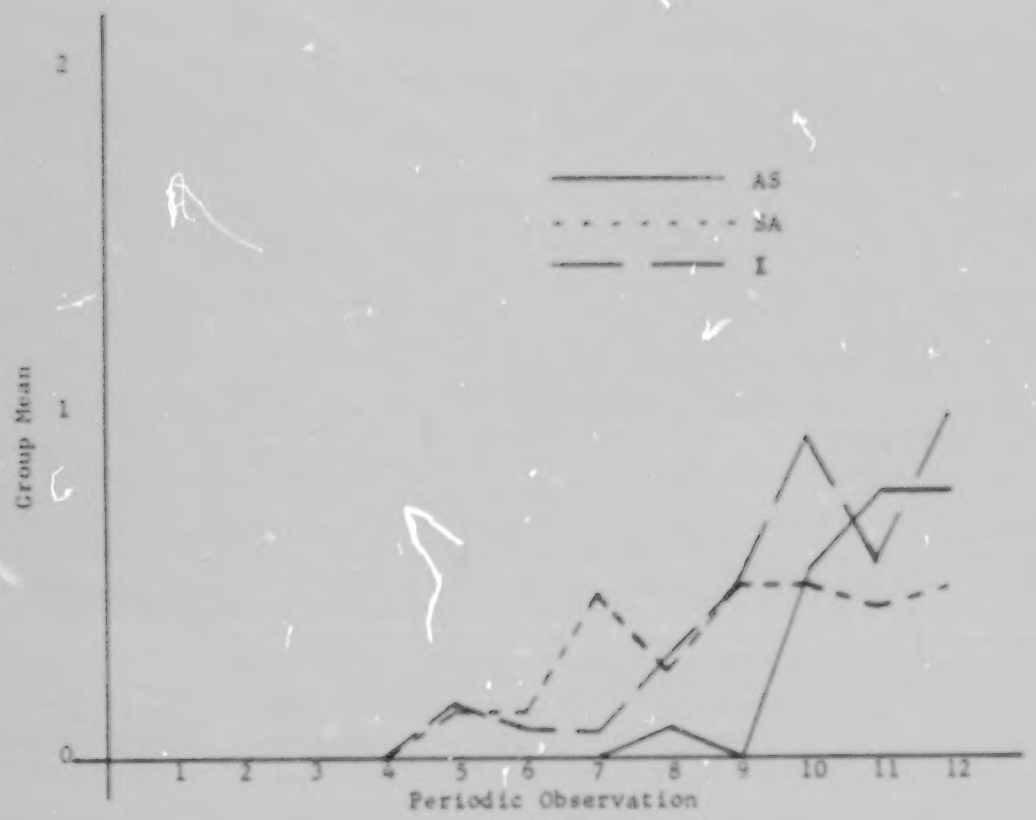
Graphs of Group Means for P1SNB Through P12SNB

Figure 3



Graphs of Group Means for P1C Through P12C

Figure 4



Graphs of Group Means for P1B Through P12B

Figure 5

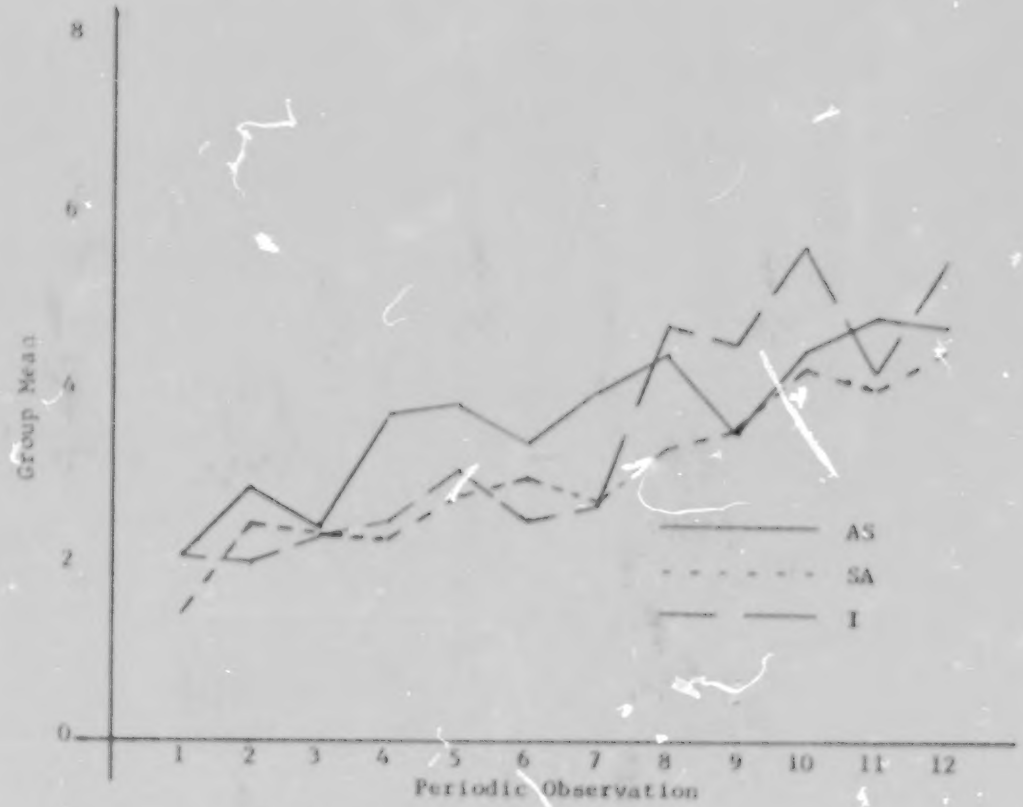
is not clearly different for any treatment group. The group means for the variables C and B, however, do appear to be different. The mean of group AS for the variable C from P4 through P8 is generally greater than that of the other two groups. Group AS has a sharp increase in performance on the variable B from P9 to P11. While the changes over time in the means of groups SA and I for the variables B and C are similar, some differences can be noted. Their means for the variable B began to increase from zero with observation P5, but the mean of group I continues to increase from P9 through P12 while that of group SA remains about the same. The means of groups SA and I for the variable C began to increase from a generally low performance level with observation P8. For the last six observations, however, the mean of group I is usually greater than that of group SA.

The total performance of each group at each periodic observation was obtained by adding the group means for the four variables ANC, SNB, C and B. Table 11 reports the group means for total performance and Figure 6 displays them. The means of the three groups for total performance were very similar for the first three observations. For the observations P4 through P7 the means of groups SA and I remain very similar while the means of group AS are noticeably greater. Beginning with observation P8, the means of group SA and I began to increase noticeably, with the mean of group I exceeding those of the other two groups at P8, P9, P10 and P12. The mean of group AS continues to increase as well, however, and group AS has the greatest

Table 11

Total Group Means for the Twelve Periodic Observations

Group	Observation											
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
AS	2.17	2.90	2.42	3.75	3.82	3.41	4.00	4.41	3.54	4.47	4.85	4.76
SA	1.43	2.44	2.42	2.33	2.80	3.00	2.71	3.38	3.56	4.26	4.06	4.44
I	2.20	2.07	2.36	2.51	3.07	2.54	2.67	4.79	4.57	5.64	4.22	5.47



Graphs of Total Performance of Groups for P1 Through P12

Figure 6

mean at observation P11. Though the means for the three groups at observations P11 and P12 do not appear to be greatly different, the mean of group SA for total performance is least for 7 of the 12 observations.

The erratic appearance of the graphs of group performance across the twelve periodic observations reflects the relatively large standard deviation of the scores. To reduce the variance of the scores, and to provide a more stable picture of group achievement during instruction, the data from the twelve periodic observations were consolidated into data for four time periods. These four time periods are referred to as observations T1, T2, T3 and T4. Scores for these observations were determined as follows: Observations T1 and T2 are observations of performance prior to the content change of treatments AS and SA that occurred immediately following P7. The scores for T1 variables accordingly are averages of the scores at P1, P2, P3 and P4, and the scores for T2 variables are averages of the scores at P5, P6 and [illegible]. Time periods T3 and T4 cover the last phase of instruction; scores for T3 variables are averages of the scores at P8 and P9 and scores for T4 variables are averages of the scores at P10, P11 and P12.

Some subjects had no scores for all of the P observations of one or more time periods. Their scores were not included in forming the estimates for any of the time periods. Therefore, the group mean estimates for the four time periods are based upon data from 15 sub-

jects from group I, 13 subjects from group AS, and 17 subjects from group SA. The estimated group means and standard errors for the variables of the observations T1 through T4 are reported in Table 12. The tests of hypothesis H1 involve the mean scores for these variables.

The variables of observations T1 through T4 are ANC, SNB, C and B and the four variables constructed from them, SUM, OP, RGP and INT. These variables have been defined earlier and are summarized in Table 6 of this chapter. The estimated group means for the variables ANC, SNB, C, B and SUM reported in Table 12 are displayed by the graphs in Figure 7. Group performances for these variables are discussed in the following paragraphs.

Group AS initially has better performance for addition (variables ANC and C) than do the other two groups; this performance stays nearly constant for the last two time periods. The subtraction performance (variables SNB and B) of group AS increased in the last two time periods though performance on variable SNB is not so drastically changed as is performance on variable B. The total performance of group AS as measured by the variable SUM is greater than that of group SA for every time period and greater than that of group I for the first two time periods.

Group SA failed to exhibit better performance on the variable SNB during T1 and T2, though their performance for variable B is somewhat better than that of groups AS and I during the time period T2. The addition performance of group SA increased during time periods T3 and

Table 12

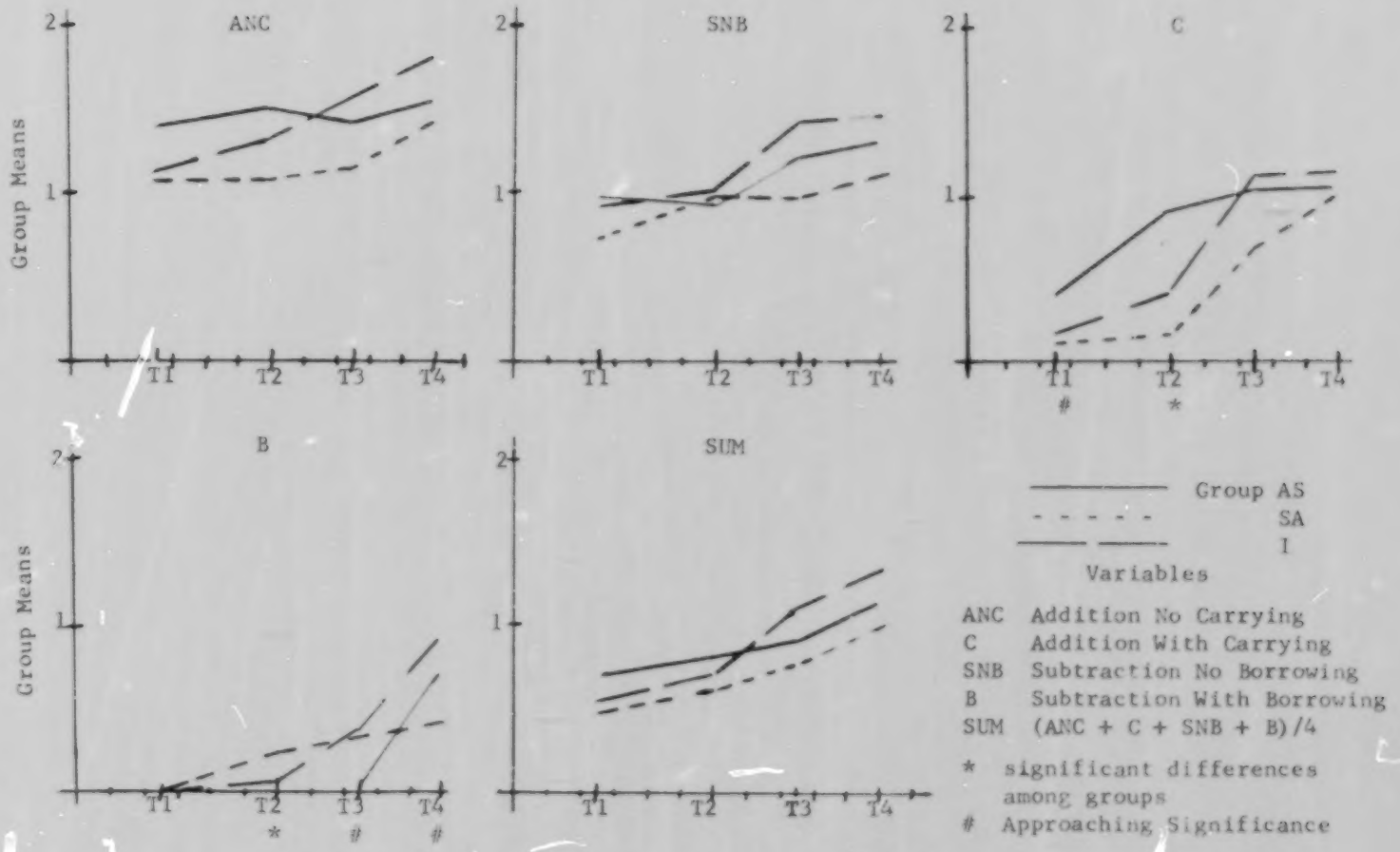
Estimated Group Means and Standard Errors for Four Time Periods

Time Period	T1			T2			T3			T4		
Group (N)	AS(13)	SA(17)	I(15)	AS(13)	SA(17)	I(15)	AS(13)	SA(19)	I(15)	AS(13)	SA(17)	I(15)
Variable												
ANC	1.40 .19	1.07 .16	1.14 .18	1.50 .19	1.09 .16	1.31 .18	1.41 .18	1.15 .15	1.56 .17	1.53 .16	1.41 .14	1.79 .15
C	.41 .10	.12 .08	.18 .10	.92 .19	.18 .16	.43 .19	1.07 .22	.70 .19	1.13 .21	1.06 .20	1.00 .17	1.15 .19
SNB	.99 .24	.72 .20	.92 .23	.95 .21	1.02 .17	1.04 .20	1.24 .18	.98 .16	1.43 .17	1.33 .21	1.12 .18	1.48 .20
B	0.- -	0.- -	0.- -	-.07 .09	.24 .07	.07 .09	0.00 .14	.34 .12	.38 .14	.71 .20	.40 .17	.93 .20
SUM	.70 .10	.48 .08	.56 .10	.82 .11	.63 .10	.71 .11	.93 .15	.79 .13	1.12 .14	1.16 .16	.99 .13	1.34 .15
OP	.41 .11	.23 .09	.20 .10	.77 .13	.01 .11	.31 .13	.62 .11	.26 .09	.44 .11	.28 .13	.44 .11	.27 .13

Table 12 (Continued)

Time Period	T1			T2			T3			T4		
Group (N)	AS(13)	SA(17)	I(15)	AS(13)	SA(17)	I(15)	AS(13)	SA(17)	I(15)	AS(13)	SA(17)	I(15)
Variable												
RGP	1.00 .21	.84 .18	.94 .20	.80 .19	.85 .16	.93 .19	.80 .14	.54 .12	.74 .14	.55 .16	.56 .14	.59 .16
INT	.00 .08	.11 .07	.02 .08	-.22 .14	.07 .12	-.04 .13	-.45 .12	-.10 .10	-.31 .11	-.08 .10	-.16 .08	.05 .10

NOTE: Entries are of the form Estimated mean
Estimated standard error



Graphs of Group Means for Four Time Periods of Instruction

Figure 7

T4 though it never quite equals that of groups AS and I for any given time period. With the exception of performance for subtraction (variables SNB and B) at T2, group SA has the least means of the three groups for all of the variables ANC, SNB, C, B and SUM for every time period.

Group I performance increased from each time period to the next with the sharpest increases occurring from T2 to T3 for variables C, SNB and B and again for variable B from T3 to T4. The result of the general increase is that all comparisons for time periods T3 and T4 favor group I.

An additional characteristic of the data for T1 through T4 is that with the exception of the performance of group SA at observation T2, the means for the four measured achievement variables are ordered from greatest to least as ANC, SNB, C, B; furthermore group means for the variable OP are never negative.

The periodic observation data that has just been described relate to hypothesis H1 which is as follows:

H1: There are no differences in group performance during instruction. This hypothesis is not directly tested, but is subdivided into four hypotheses concerning the group means for the variables ANC, SNB, C and B. These hypotheses are as follows:

H1.1 There are no differences among the group means for the variable ANC at each of the four time periods.

H1.2 There are no differences among the group means for the variable SNB at each of the four time periods.

H1.3 There are no differences among the group means for the variable C at each of the four time periods.

H1.4 There are no differences among the group means for the variable B at each of the time periods T2, T3 and T4.

Since no subject had a nonzero score for T1B, the observation T1 is not included in the statement of H1.4.

Each of the hypotheses H1.1 through H1.4 was tested by the multivariate analysis of variance F test of group differences using the means and standard errors contained in Table 12. The statistics for the F tests for each of these hypotheses are presented in Table 13.

The F ratios for the hypotheses H1.1 and H1.2 are 1.14 and .66, respectively. These values are not significant, and thus these hypotheses are not rejected. The F ratio for hypothesis H1.3 is 1.65; since p is less than .14 this value is approaching significance. The F ratio for H1.4 is 2.43; this is significant at the .05 level and so H1.4 is rejected. Since H1.4 is rejected and the test for H1.3 is approaching significance, the hypothesis H1 is judged to be false and diffferences among the groups with respect to the variables B and C, borrowing and carrying, are assumed to exist.

Further specification of differences is desirable. Accordingly, F tests of group differences for the variables B and C of each time period were made. These are reported in Table 13. The univariate F ratios for T2C and T2B are 4.18 and 3.39, respectively; these are significant. The F ratios for T1C, T3B, and T4B are approaching significance ($p < .15$). The conclusion of this aspect seems to be

Table 13

MANOVAs and ANOVAs of Group Means for T Variables

Multivariate Tests

Variables	F	p	df
T1ANC, T2ANC, T3ANC, T4ANC	1.14	.36	8/44
T1SNB, T2SNB, T3SNB, T4SNB	.66	n.s.	8/44
T1C, T2C, T3C, T4C	1.65	.14#	8/44
T2B, T3B, T4B	2.43	.05*	6/46
T1SUM, T2SUM, T3SUM, T4SUM	1.30	.27	8/44
T1OP T2OP, T3OP, T4OP	2.64	.02*	8/44
T1RGP, T2RGP, T3RGP, T4RGP	.37	n.s.	8/44
T1INT, T2INT, T3INT, T4INT	1.23	.31	8/44

Univariate Tests

	Hyp. Mean Squares			df
T1C	.29	2.46	.11#	2/25
T2C	1.80	4.18	.03*	2/25
T3C	.77	1.36	.28	2/25
T4C	.08	.18	n.s.	2/25
T2B	.31	3.39	.05*	2/25
T3B	.54	2.25	.13#	2/25
T4B	1.02	2.14	.14#	2/25
T1OP	.16	1.22	.32	2/25
T2OP	1.87	9.47	.001*	2/25
T3OP	.41	2.88	.08*	2/25
T4OP	.14	.69	n.s.	2/25

*Significant

#Approached significance

that the group differences are restricted to variables B and C and are greatest for the time period T2. There is some further evidence here for additional group differences for the variables B at T3 and T4, and for the variable C and T1.

The meaning of these differences can be seen from the graphs of Figure 7 for the variables C and B. The mean scores of groups SA and I for T1C and T2C are relatively close together and less than the mean of group AS. The mean scores of groups AS and I for T2B are practically zero, while the mean of group SA is somewhat greater than zero. Any differences on T3B favor groups SA and I over group AS, and any differences on T4B favor groups I and AS over group SA.

A further search for the meaning of group differences was conducted; this involved looking for differences in the group means on the constructed variables. Four supplementary hypotheses concerning these constructed variables were tested. The supplementary hypotheses are:

H1.5 There are no differences among the group means for the variable SUM at each of the four time periods.

H1.6 There are no differences among the group means for the variable OP at each of the four time periods.

H1.7 There are no differences among the group means for the variable RGP at each of the four time periods.

H1.8 There are no differences among the group means for the variable INT at each of the four time periods.

The statistics for the multivariate tests of these four hypotheses are also found in Table 13. The F ratio for hypothesis H1.6 is 2.64 and significant ($p < .02$); the other three F ratios are not significant. Hypothesis H1.6 is therefore rejected and univariate F tests of group differences on T1OP through T4OP were consequently made. The statistics for these univariate tests are also reported in Table 13. These indicate significant group differences on T2OP ($p < .001$) and on T3OP ($p < .08$).

The means of the three groups on T2OP are .77, .01, and .31 for groups AS, SA, and I, respectively. The order of these means corresponds with the differences in instruction during the first phase. An interesting fact is that the mean of group SA is not negative. The means on T3OP are .62, .26, and .44 for groups AS, SA, and I. The order of these means also corresponds to the instruction of the first phase.

Summary of the Periodic Observation Data

During the first phase of instruction, covered by T1 and T2, the instruction of group AS covered all addition activities, the instruction of group SA covered all subtraction activities, and the instruction of group I was concerned with manipulative models for both addition and subtraction and with expanded notation algorithms. During the last phase of instruction, which began immediately following the second time period, group AS studied subtraction, group SA studied addition, and group I made the transition from expanded notation

algorithms for addition and subtraction to compact algorithms for the same two operations.

As was anticipated the group means for the variables ANC and SNB were greater than those for C and B, respectively, at every observation. Also as anticipated the group means for ANC and C were greater than those for SNB and B, respectively, at the end of instruction. However, the following characteristics of the data were contrary to expectations. (1) The achievement profiles of the three groups for the variables ANC and SNB were similar in appearance; there were not significant differences among these group means for any of the four time periods. (2) The means for the variables ANC and SNB were relatively large for the first time period, and they did not change sharply from one time period to the next. (3) The means of each group for the variables ANC, SNB, C, and B for every time period, with a single exception, were ordered from largest to smallest as ANC, SNB, C, and B. The only departure from this pattern was for the time period T2 where the means for group SA were ordered from largest to smallest as ANC, SNB, B, and C. (4) The mean score of every group for the variable OP was never less than zero at any observation.

The performance profiles of the groups for the variables B and C did show marked differences, however, in the manner that would be expected from the instructional treatments. Significant differences were found for the variables C, B, and OP of the second time period,

and there was some evidence of group differences for these same variables during the third time period. The group means for the variable C at T2 were ordered from greatest to least as AS, I, and SA, and at T3 as I, AS, and SA. The group means for the variable B at T2 were ordered from greatest to least as SA, I, and AS, and at T3 as I, SA, and AS. The order of the group means for the variable OP at both T2 and T3 was AS, I, and SA.

Though the differences were not found to be significant, the group means for total performance during the instruction (the variable SUM) were ordered as follows (see Figure 7). The mean of group AS exceeded that of the other two groups during the first two time periods; the mean of group I exceeded that of the other two groups for the last two time periods and the mean of group SA was the least for all four time periods.

Sequence of Data Analyses for 02 and 03 Data

The data from the tests of 02 and 03 relate to achievement levels after instruction and retention two weeks later, respectively. The 03 - 02 data are used as direct measures of forgetting. Since the tests administered at 02 were also administered at 03, and since the 03 - 02 data were derived from the 02 and 03 data in the obvious way, the procedures followed in the analysis of the 02, 03, and 03 - 02 data were the same.

Most of the statistical tests of group differences involved analysis of covariance. Therefore, the first step in the overall plan of the analysis was to determine covariates for the analysis;

the next section describes the way the covariates were selected. After the selection of the covariates, the data from each observation were analyzed separately. For each observation the data for the achievement variables, the total power test variables, and the difference power test variables were also analyzed separately. A multivariate F test was first used to test the significance of the overall group differences for each set of variables. If the differences among the groups were significant or approached significance, three further steps were taken to identify dimensions and factors of the overall group differences.

The first of these descriptive steps was a discriminant analysis of the multivariate space (Bock, 1966); Bartlett's chi square approximation (Bartlett, 1947) to the maximum likelihood criterion was used to test the significance of the canonical variates or discriminant functions produced by this analysis. When a function was found to be significant, the weights assigned to the dependent variables were used to identify variables or variable contrasts that contributed most to the differences. The function values for each group were used to identify overall similarities or dissimilarities among the groups. The second step in describing group differences was to test the significance of the differences among the means of the three groups for each individual variable, and the last step was to perform multivariate and univariate tests of the differences between pairs of groups.

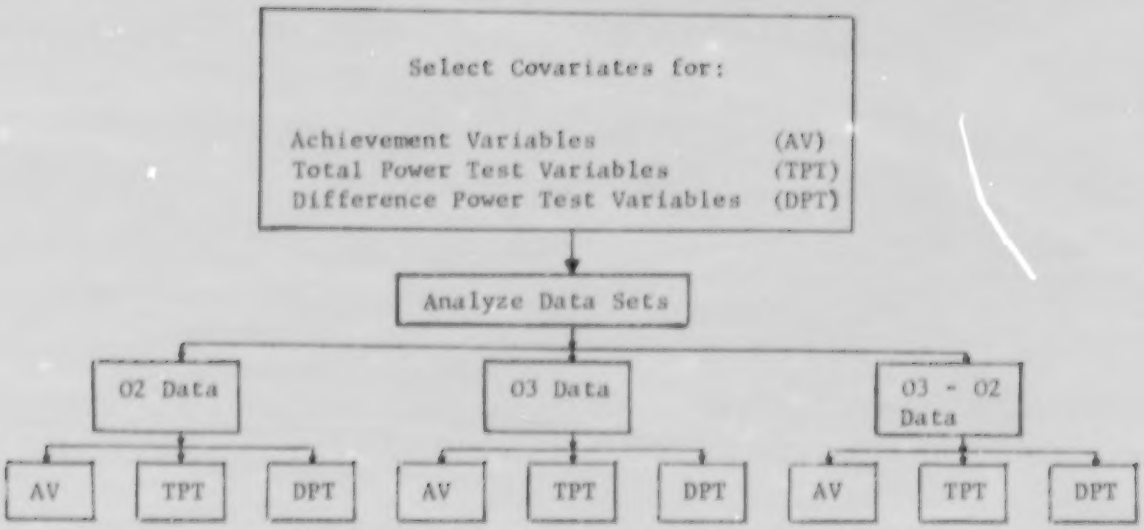
If the multivariate test of overall group differences was not significant, the multivariate and univariate tests of the second and third steps listed above were still executed. However, the purpose was not to describe differences inferred to exist, but rather to describe possible trends in the data and to provide information that may lead to further research.

The plan followed in reporting the data derived from 02 and 03 is as follows. The selection of the covariates used in the analysis is reported first followed by the data from 02, 03, and 03 - 02 in that order. There are three variable sets for each observation: the achievement variables, the total power test variables, and the difference power test variables. For each variable set the reporting procedure is to present the estimated group means, then the adjusted group means (if covariates were used), then the test of overall group differences, and finally present the characterization of any overall group differences.

The entire analysis is summarized by Figures 8A and 8B.

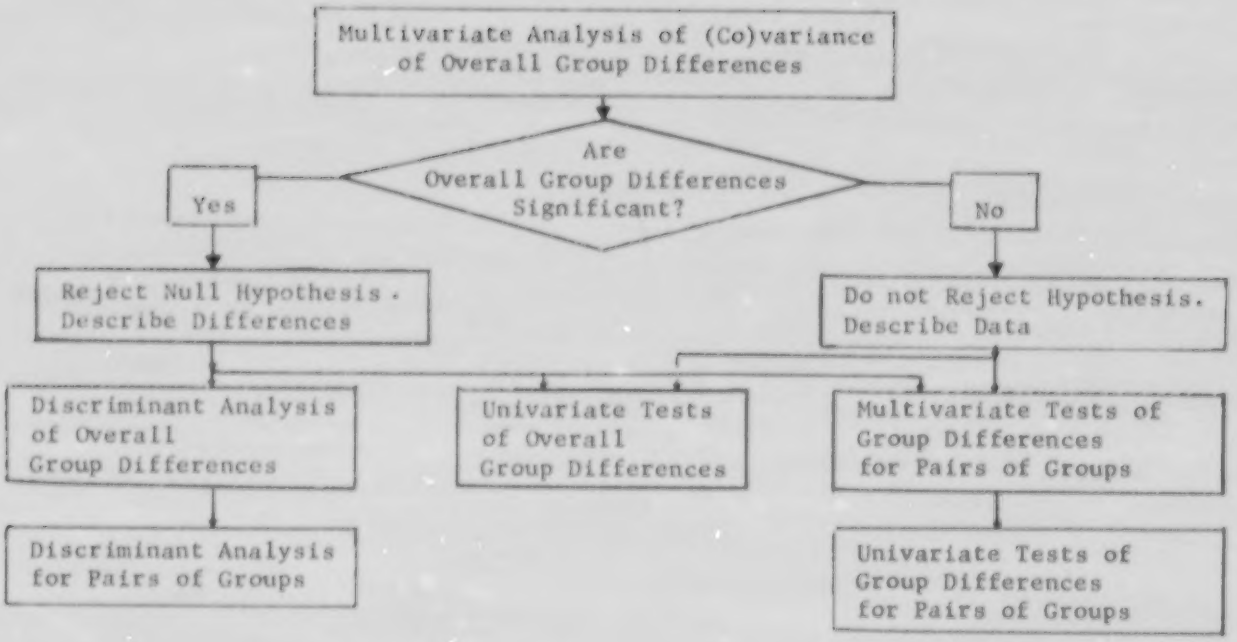
The Selection of Covariates

Since the observations 02, 03, and 03 - 02 all involve three categories of variables; as have just been outlined (see Figure 8A), covariates for a category were selected independently of the selection of the covariates for either of the other two categories. The covariates were selected from the 01 variables and the status variables that were identified earlier in this chapter (see Table 6). The data



Analysis of Post Test Data

Figure 8A


Multivariate Analysis of (Co)variance of Overall Group Differences
Are Overall Group Differences Significant?
Yes
No
Reject Null Hypothesis. Describe Differences
Do not Reject Hypothesis. Describe Data
Discriminant Analysis of Overall Group Differences
Univariate Tests of Overall Group Differences
Multivariate Tests of Group Differences for Pairs of Groups
Discriminant Analysis for Pairs of Groups
Univariate Tests of Group Differences for Pairs of Groups

Hypothesis Testing for Each Branch of the Post Test Data

Figure 8B

of the 41 subjects having complete 02 and 03 data, and that of another subject with only 03 data, were used in the selection process.

The 02 and 03 variables were grouped into three sets of dependent variables, one set for each of the three categories of dependent variables. The first set consisted of the achievement variables +F, -F, ANC, C, SNB, and B for both 02 and 03. The second set consisted of the total power test variables SUMB1 through SUMB7 for both observations 02 and 03, and the third set consisted of the difference power test variables DIFB1 through DIFB7 for both 02 and 03. The set of potential covariates consisted of the status variables MET1, MET2, and MET3, and the 01 variables +F, -F, ANC, and SNB.

For each of the three categories the first covariate was chosen in the following way. The scores for each of the potential covariates were entered one at a time into the randomized block model as independent variables against the set of dependent variables of each category. The multiple correlations of each of the independent variables with each set of dependent variables were found, and the strength of each association was tested by an F test. If any variable among the set of independent variables accounted for a significant amount of the total variance in the multivariate space of a category of dependent variables, that variable with the greatest multiple correlation was chosen at the first covariate for that category. If none of the F ratios were significant ($p < .10$) then no covariates were chosen for that category of dependent variables.

When one covariate was determined for a category of dependent variables, an attempt was made to determine a second. The multiple correlations of each pair of independent variables that included the first selected covariate, with the set of dependent variables were formed. The amount of additional variance accounted for by the inclusion of the second independent variable was then tested. If a second independent variable accounted for a significant amount of additional variance in the multivariate space of the dependent variables ($p < .10$), the variable that accounted for the greatest amount of additional variance was taken as a second covariate. If no other independent variable accounted for a significant amount of additional variance, the search for additional covariates was terminated.

If a second covariate were determined, the search for additional covariates was to be carried out in a manner analogous to that for the second with an additional step. If three or more covariates could have been chosen, it would have been necessary to determine which subset, that did not contain one of them, related most strongly to the dependent variables; then test to see if the remaining covariate accounted for a significant amount of additional variance that the subset did not. This could result in the dropping of a covariate previously selected and a new search for an additional covariate. However, such procedures never became necessary as in no case was even a third covariate determined. The following paragraphs report

the covariates selected first for the achievement variables, then for the total power test variables, and finally for the difference power test variables.

The Covariates for the Achievement Variables

Table 14 contains the statistics relating to the selection of the covariates for the analysis of the achievement variables. The variables MET1 and 01SNB were found to be significantly related to the set of 12 dependent variables +F, -F, ANC, C, SNB, and B for both 02 and 03. Since the relationship was strongest for MET1 ($p < .01$), it was taken as the first covariate. After the group means were adjusted by the variable MET1, the amount of additional variance accounted for by the variable 01SNB was found to be significant ($p < .09$), and so 01SNB was the second covariate chosen. No other potential covariates were found to be significantly related to the dependent variables.

The Covariates for the Power Test Variables

Table 15 contains the statistics relating to the selection of the covariates for the total power test variables. The variable MET2 was the only variable found to be significantly related to the dependent variables 02SUMB1 through 02SUMB7 and 03SUMB1 through 03SUMB7 ($p < .01$). No remaining potential covariate was found to be significantly related to the dependent variables after the means were adjusted for MET2.

The statistics relating to the selection of covariates for the difference power test variables are also found in Table 15. No available

Table 14

F Tests of No Association Between the Achievement Variables of 02 and 03 and the Potential Covariates

Covariates Eliminated	Independent Variables	F Ratio	p	df Numerator/Denominator
None	MET1	6.82	.01	12/10
None	MET2	2.10	.13	12/10
None	MET3	1.96	.15	12/10
None	01+F	1.62	.23	12/10
None	01-F	2.16	.12	12/10
None	01ANC	2.24	.11	12/10
None	01SNB	3.31	.04	12/10
MET1	MET2	1.27	.37	12/9
MET1	MET3	.47	.89	12/9
MET1	01+F	.55	.84	12/9
MET1	01-F	.92	.57	12/9
MET1	01ANC	1.47	.29	12/9
MET1	01SNB	2.49	.09	12/9
MET1,01SNB	MET2	1.14	.44	12/8
MET1,01SNB	MET3	.45	n.s.	12/8
MET1,01SNB	01+F	.42	n.s.	12/8
MET1,01SNB	01-F	1.05	.49	12/8
MET1,01SNB	01ANC	.83	n.s.	12/8
None	MET1,01SNB	3.97	.01	24/18

Table 15

F Tests of No Association Between the Power Test Variables of 02 and 03 and the Potential Covariates

(Tests for the 02 and 03 Sum Power Test Variables)

Covariates Eliminated	Independent Variables	F Ratio	p	df Numerator/Denominator
None	MET1	2.21	.14	14/8
None	MET2	5.85	.01	14/8
None	MET3	1.42	.32	14/8
None	01+F	.87	n.s.	14/8
None	01-F	1.75	.22	14/8
None	01ANC	1.23	.40	14/8
None	01SNB	.56	n.s.	14/8
MET2	MET1	1.64	.26	14/7
MET2	MET3	.81	n.s.	14/7
MET2	01+F	1.37	.36	14/7
MET2	01-F	2.12	.16	14/7
MET2	01ANC	1.70	.25	14/7
MET2	01SNB	.53	n.s.	14/7

(Tests for the 02 and 03 Difference Power Test Variables)

None	MET1	1.00	.53	14/8
None	MET2	.26	n.s.	14/8
None	MET3	1.21	.41	14/8
None	01+F	.49	n.s.	14/8
None	01-F	.26	n.s.	14/8
None	01ANC	1.03	.51	14/8
None	01SNB	.70	n.s.	14/8

covariates were found to be significantly related to the 14 power test variables 02DIFB1 through 02DIFB7 and 03DIFB1 through 03DIFB7.

The Post Test Data from Observation 02

The data from the observation 02 relates to group performance immediately following instruction. The group means and standard errors for the 02 variables were estimated using the randomized block design and the scores of 13 subjects from group AS, 16 subjects from group SA and 15 subjects from group I; the estimated group means and standard errors are contained in Table 14.

The means and the analyses relating to them are discussed in the following order; the achievement variables are considered first, then the total power test variables, and finally the difference power test variables.

The 02 Achievement Variables

The 02 achievement variables were measures of student performance under timed conditions. There are six basic achievement variables, +F, -F, ANC, C, SNB, and B and four constructed variables, SUM, OP, RGP, and INT, that enter into the description of group performance following instruction.

The estimated group means for the 02 achievement variables are presented in Table 16. The null hypothesis to which these data apply is as follows.

H2,02 There are no differences among the group means for the achievement measures of 02.

Table 16

Estimated Group Means and Standard Errors for 02 Variables

Variable	AS (13)		SA (16)		I (15)	
	Estimated Mean	Estimated Standard Error	Estimated Mean	Estimated Standard Error	Estimated Mean	Estimated Standard Error
+F	8.23	.97	4.71	.84	4.96	.88
-F	6.68	1.00	3.69	.86	3.80	.91
ANC	2.49	.34	2.22	.29	2.47	.31
C	4.11	.80	3.11	.69	4.11	.73
SNB	2.59	.38	1.88	.33	2.34	.35
B	3.43	.84	.75	.72	2.61	.75
SUM	2.21	.35	1.51	.30	2.04	.32
OP	.12	.19	.76	.16	.44	.17
RGP	.66	.21	1.09	.18	.73	.19
INT	-.22	.17	-.42	.15	-.31	.16
SUMB1	1.65	.55	1.02	.47	1.02	.50
SUMB2	5.44	.76	5.06	.65	5.06	.69
SUMB3	3.07	.82	3.13	.70	3.84	.74
SUMB4	7.28	.52	7.62	.44	6.77	.47
SUMB5	6.16	.53	6.54	.45	6.15	.48
SUMB6	4.18	.76	3.28	.65	4.92	.69
SUMB7	5.93	.79	6.80	.67	5.73	.71
DIFB1	-.44	.26	-.05	.23	.17	.24
DIFB2	-.14	-.30	.91	.26	1.05	.27
DIFB3	-.29	-.36	1.30	.30	.45	.32
DIFB4	-.26	.12	.01	.10	.15	.11
DIFB5	.27	.24	.52	.20	.48	.22
DIFB6	.07	-.54	1.24	.47	.66	.49
DIFB7	.25	-.20	.61	.18	.47	.19

This hypothesis was tested by multivariate analysis of covariance on the mean scores for the variables +F, -F, ANC, C, SNB, and B using the covariates MET1 and C1SNB. This test is reported following the description of the group means.

The group means for the achievement variables of 02 adjusted by the covariates MET1 and O1SNB are listed in Table 17. The adjusted means of group AS differed little from the estimated means of Table 16, while the means of group SA were increased and those of group I were decreased. Though these adjustments did change the order of the group means for the variables +F, -F, and C, these changes do not appear to be large.

The adjusted means of group AS exceeded those of the other groups for all six of the basic 02 achievement variables, +F, -F, ANC, C, SNB, and B, while the means of group SA were the least for the variables ANC, SNB, C, and B, and the means of group I were least for the variables +F and -F.

The group means for the constructed 02 achievement variables SUM, OP, RGP, and INT summarized the information from ANC, C, SNB, and B in the following way. The total performance as measured by the variable SUM was greatest for group AS and least for group SA. While all the estimated group means for the variable OP were positive, group AS had the least mean and group SA the greatest; hence, the difference in performance for addition and subtraction was least for group AS and greatest for group SA.

All of the group means for the variable RGP were positive; group AS had the least mean and group SA the greatest. Hence, the difference in performance for non-regrouping and regrouping items was least for group AS and greatest for group SA.

The group means for the variable INT were negative, showing that the difference in performance between the non-regrouping and regrouping items for subtraction was greater than the corresponding difference for addition. The mean of group AS was least in absolute value and the mean of group SA was greatest in absolute value; hence, group AS had the least interaction between operation and the need to regroup, and group SA had the greatest.

The statistics for the multivariate F test of hypothesis H2.02 are given in Table 18. The overall F ratio is significant ($p < .03$) and the hypothesis is rejected. Therefore, the analyses as outlined earlier in Figure 8B were undertaken to specify the differences.

Discriminant analysis for 02 achievement variables. The results of the discriminant analyses of the group means for the 02 variables +F, -F, ANC, C, SNB, and B are given in Table 19. The first canonical variate for the means of all three groups accounted for 89.2% of the canonical variation, and the second the remaining 10.8%. While the contrasts taken together provided a significant contrast among the groups ($p < .03$), the second contrast by itself did not ($p < .58$). The first canonical variate is therefore taken as the best single descriptor of the multivariate group differences.

TABLE 17

Adjusted Group Means for the 02 Achievement Variables
Covariates: MET1, 01SNB

	Variables									
Groups(N)	+F	-F	ANC	C	SNB	B	SUM	OP	RGP	INT
AS(13)	8.24	6.55	2.48	4.09	2.59	3.39	2.20	.12	.66	-.23
SA(16)	5.09	4.07	2.34	3.37	2.01	1.07	1.64	.74	1.06	-.41
I(15)	4.56	3.55	2.36	3.87	2.21	2.31	1.92	.46	.74	-.31

TABLE 18

Overall Groups MANCOVA and ANCOVAs for 02 Achievement Variables
Covariates: MET1, 01SNB

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	multivariate	2.39	.03*	12/34 N/D
+F	47.16	11.13	.001*	2/22 hyp/error
-F	30.60	10.62	.001*	2/22
ANC	.06	.10	n.s.	2/22
C	1.73	.39	n.s.	2/22
SNB	1.08	1.12	.35	2/22
B	16.81	5.10	.02*	2/22
SUM	.97	1.72	.21	2/22
OP	1.21	2.85	.08*	2/22
RGP	.58	1.12	.35	2/22
INT	.09	.28	n.s.	2/22

* Significant

Table 19

Discriminant Analysis: 02 Achievement Test Variables Adjusted Means

Covariates: MET1, O1SNB

First Canonical Variate - 3 Groups - 89.2% of Canonical Variation

	Discriminant Function Coefficients for Variables						Function Values for Groups		
	+F	-F	ANC	C	SNB	B	AS	SA	I
Raw	-.2435	-.31745	.8904	.0430	-.8076	-.1589	-4.33	-2.10	-2.12
Standardized	-.5011	-.5389	.7177	.0904	-.7897	-.2385	--	--	--

Second Canonical Variate - 2 Groups - 10.8% of Canonical Variate

Raw	-.0263	-.2875	-.3398	.0507	-.0889	.5433	-1.12	-1.53	-.69
Standardized	-.0541	-.4879	-.2739	.1067	-.0869	.9865	--	--	--

I vs. AS Canonical Variate

Raw	-.2440	-.3666	.8090	.0519	-.8097	-.0521	-4.47	--	-2.21
Standardized	-.5022	-.6222	.6521	.1092	-.7917	-.0945	--	--	--

I vs. SA Canonical Variate

Raw	-.0181	-.2767	-.3693	.0492	-.0619	.5483	--	-1.45	-.62
Standardized	-.0373	-.4697	-.2976	.1036	-.0606	.9355	--	--	--

Table 19 (Continued)

AS vs. SA Canonical Variate

Discriminant Function Coefficients for Variables							Function Values for Groups		
	+F	-F	ANC	C	SNB	B	AS	SA	I
Raw	-.2349	-.2611	.9367	.0333	-.7787	-.2533	-4.06	-1.79	--
Standardized	-.4834	-.4431	.7550	.0699	-.7615	-.4599	--	--	--

Bartlett's Chi Square Test of Significance

Variate(s)	χ^2	df	p
First and Second-3 Groups	23.89	12	.03*
Second-3 Groups	3.84	5	.58
I vs. AS	16.79	6	.02*
I vs. SA	3.75	6	.72
AS vs. SA	17.10	6	.01*

* Significant

The magnitudes of the standardized coefficients of the first discriminant function indicated the variables contributing most to the differences listed in order from most to least were SNB, ANC, -F, and +F. The variables B and C were relatively unimportant in this contrast. The coefficients for the variables ANC and C were positive while the coefficients for the variables SNB, B, +F, and -F were negative. Hence, the group differences were greatest when addition performance was contrasted with both subtraction performance and facts performance. The values of the first function for the groups AS, SA, and I were -4.33, -2.10, and -2.12, respectively; hence, the contrast was primarily between group AS and both of the other groups.

The discriminant analysis for each pair of groups provided no new information. The canonical variate for the pair of groups I and SA was not significant, while the functions for the pairs I and AS ($p < .02$) and AS and SA ($p < .01$) were. The two significant variates for the pairs of groups were very similar to the first canonical variate for all three groups.

Overall groups univariate analysis - 02 achievement variables. The next step in the description of the differences among the three groups was the testing of the significance of the differences among the three groups for each of the 02 achievement variables. The statistics for the F tests of these differences are given in Table 18. Significant differences existed among the groups for the variables +F ($p < .001$), -F ($p < .001$), B ($p < .02$), and OP ($p < .08$).

Pairs of groups analysis - 02 achievement variables. The multivariate and univariate differences between pairs of groups for the 02 achievement variables were also tested for significance. The statistics relating to these tests are reported in Table 20. The multivariate differences between the groups I and SA were not significant ($F = .62$), however, the multivariate differences between groups AS and SA were significant ($p < .01$), as were the differences between I and AS ($p < .02$). The univariate differences between groups AS and SA were significant for the 02 variables +F, -F, B, SUM, and OP, and the univariate differences between groups I and AS were significant only for +F and -F. While the overall differences between groups I and SA were not significant, the difference for the variable B was significant.

Summary of 02 Achievement Variable Comparisons

The multivariate differences among the three groups for the 02 achievement variables +F, -F, ANC, C, SNB, and B were significant and the hypothesis H2.02 is rejected. The discriminant analysis indicated that the differences were primarily between group AS and the other two groups. All of these differences favored group AS. The differences among the groups were maximized by contrasting their performance for addition with that for subtraction and the addition and subtraction facts.

When the groups were compared by pairs, the multivariate differences between groups I and SA were not significant while the overall differences were significant for the groups AS and SA, and I and AS.

The univariate analyses identified significant differences for the variables +F, -F, B, SUM, and OP. The means of group AS were significantly greater than those of group SA for the variables +F, -F, B, SUM, and were significantly greater than those of group I for +F and -F. The mean of group AS for the variable OP was significantly less than that of group SA. The mean of group I was significantly greater than that of group SA for the variable B.

The 02 Total Power Test Variables

The 02 total power test variables SUMB1 through SUMB7 were measures of total power test performance following instruction. The estimated group means for these variables are reported in Table 16. The null hypothesis to which these data apply is as follows.

H3.02 There are no differences among the group means for the power test measures of total performance at 02.

This hypothesis was tested by multivariate analysis of covariance of the mean scores for the 02 variables SUMB1 through SUMB7 using the covariate MET2. This test is reported following the description of the group means.

The estimated group means for the 02 total power test variables adjusted by the covariate MET2 are listed in Table 21. The general effect of the adjustment was to decrease the means of group AS slightly and to increase those of groups SA and I by very small amounts. The rank orders of the adjusted group means were the same as for the estimated group means before the scores for MET2 were entered.

Table 20

Pairs of Groups MANCOVAs and ANOVAs for 02 Achievement Variables
Covariates: MET1, 01SNB

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	4.02	.02*	6/17 N/D
+F	82.84	19.55	.001*	1/22 hyp/error
-F	55.06	19.11	.001*	1/22
ANC	.08	.12	n.s.	1/22
C	.31	.07	n.s.	1/22
SNB	.88	.93	n.s.	1/22
B	7.15	2.17	.16	1/22
SUM	.50	.89	n.s.	1/22
OP	.74	1.74	.21	1/22
RGP	.04	.08	n.s.	1/22
INT	.04	.12	n.s.	1/22

I vs. SA

+F, -F, ANC, C, SNB, B	(multivariate)	.62	n.s.	6/17 N/D
+F	1.89	.45	n.s.	1/22 hyp/error
-F	1.88	.65	n.s.	1/22
ANC	.00	.01	n.s.	1/22
C	1.67	.38	n.s.	1/22
SNB	.29	.30	n.s.	1/22
B	10.49	3.18	.09*	1/22
SUM	.51	.91	n.s.	1/22
OP	.53	1.24	.28	1/22
RGP	.70	1.34	.26	1/22
INT	.06	.18	n.s.	1/22

Table 20 (Continued)

AS vs. SA

+F, -F, ANC, C, SNB, B	(multivariate)	4.13	.01*	6/17 N/D
+F	61.24	14.45	.001*	1/22 hyp/error
-F	37.72	13.09	.002*	1/22
ANC	.11	.17	n.s.	1/22
C	3.20	.72	n.s.	1/22
SNB	2.12	2.22	.16	1/22
B	33.27	10.09	.005*	1/22
SUM	1.94	3.43	.08*	1/22
OP	2.42	5.70	.03*	1/22
RGP	.99	1.91	.18	1/22
INT	.19	.56	n.s.	1/22

* Significant

Table 21

Adjusted Group Means for 02 Total Power

Test Variables Covariate: MET2

Groups(N)	Variables						
	SUMB1	SUMB2	SUMB3	SUMB4	SUMB5	SUMB6	SUMB7
AS(13)	1.50	5.32	2.93	7.21	6.10	3.98	5.81
SA(16)	1.10	5.13	3.21	7.66	6.57	3.39	6.87
I(15)	1.09	5.12	3.91	6.79	6.18	5.01	5.79

There was no consistent rank order for the adjusted group means. While most of the differences do not appear to be large, the mean of group I for the variable SUMB6 was about two and one half standard error units greater than that of group SA.

The statistics for the multivariate F test of hypothesis H3.02 are given in Table 22. The overall F was .94, and hence the hypothesis is not rejected. The subsequent tests of group differences, as outlined earlier in Figure 8B, were made to discover any trends not indicated by the overall multivariate test.

The statistics relating to the overall groups univariate tests are reported in Table 22, and those for the tests of the differences between pairs of groups are reported in Table 23. The univariate test of overall group differences for the variable SUMB6 approached significance ($p < .15$), and the univariate test of the difference between group I and SA for this variable was significant ($p < .06$). All other multivariate and univariate tests relating to the group means for these seven power test variables were not significant.

The O2 Difference Power Test Variables

The O2 difference power test variable DIFB1 through DIFB7 were power test measures of differences in performance for similar addition and subtraction behaviors following instruction. Though these variables have been described earlier (see Table 6), it should be recalled here that a negative score indicates for a pair of similar behaviors the subtraction score was greater than the addition score.

Table 22

Overall Groups MANCOVA and ANCOVA of O2 Total Power Test Variables
Covariate: MET2

Variables	Hypothesis Mean Squares	F	p	df
SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.94	n.s.	14/34 N/D
SUMB1	.65	.27	n.s.	2/23 hyp/error
SUMB2	.16	.03	n.s.	2/23
SUMB3	3.32	.50	n.s.	2/23
SUMB4	2.60	.90	n.s.	2/23
SUMB5	.84	.27	n.s.	2/23
SUMB6	9.45	2.12	.15	2/23
SUMB7	5.14	.81	n.s.	2/23

Table 23

Pairs of Groups MANCOVA and ANCOVA for O2 Total Power Test Variables
Covariates MET2

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.98	n.s.	7/17 N/D
SUMB1	1.05	.43	n.s.	1/23 hyp/error
SUMB2	.26	.04	n.s.	1/23
SUMB3	5.99	.90	n.s.	1/23
SUMB4	1.09	.37	n.s.	1/23
SUMB5	.03	.01	n.s.	1/23
SUMB6	6.62	1.48	n.s.	1/23
SUMB7	.00	.00	n.s.	1/23

I vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	1.15	.38	7/17 N/D
SUMB1	.00	.00	n.s.	1/23 hyp/error
SUMB2	.00	.00	n.s.	1/23
SUMB3	3.45	.52	n.s.	1/23
SUMB4	5.21	1.79	.20	1.23
SUMB5	1.11	.36	n.s.	1/23
SUMB6	18.50	4.14	.06*	1/23
SUMB7	8.19	1.28	.27	1/23

Table 23 (Continued)

AS vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.71	n.s.	7/17 N/D
SUMB1	.99	.40	n.s.	1/23 hyp/error
SUMB2	.23	.04	n.s.	1/23
SUMB3	.50	.07	n.s.	1/23
SUMB4	1.20	.41	n.s.	1/23
SUMB5	1.36	.44	n.s.	1/23
SUMB6	2.13	.48	n.s.	1/23
SUMB7	6.86	1.08	.31	1/23

* Significant

The estimated group means for the 02 difference power test variables are found in Table 16. The null hypothesis to which these data apply is as follows.

H4.02 There are no differences among the group means for the 02 power test measures of differential performance for addition and subtraction.

Since no covariates were entered into the analysis model for estimating the means for the variables DIFB1 through DIFB7, hypothesis H4.02 was tested by multivariate analysis of variance of the estimated group means for the 02 variables DIFB1 through DIFB7. This test is reported following the description of the group means.

The means of group AS were less than those of both groups SA and I for all of the variables DIFB1 through DIFB7, and the scores of group AS for the variables DIFB1 through DIFB4 were negative. The means for groups SA and I for these seven variables were positive with one exception; the mean of group SA for the variable DIFB1 was -.05. These data indicate then that addition performance was generally better than subtraction performance for both groups SA and I, but that the converse may have been true for group AS. There are no apparent patterns of differences between the means of groups SA and I for the power test variables DIFB1 through DIFB7.

The statistics relating to the test of hypothesis H4.02 are reported in Table 24. The multivariate F test for this hypothesis approached significance ($p < .11$), and although H4.02 cannot be rejected on the basis of this test, it does provide sufficient reason for the succeeding steps, assuming real differences were present as outlined in Figure 8B.

Discriminant analysis for O2 difference power test variables. The results of the discriminant analyses of the group means for the O2 variables DIFB1 through DIFB7 are presented in Table 25. The first canonical variate for the means of all three groups accounted for 82.5% of the canonical variation, and the second the remaining 17.5%. While the contrasts taken together provided a contrast among the groups was significant ($p < .10$), the second variate by itself did not ($p < .58$). Hence, only the first is considered further.

Table 24

Overall Groups MANOVA and ANOVAs for O2 Difference Power Test Variables

Variables	Hypothesis Mean Squares	F	p	df
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	1.69	.11#	14/36 N/D
DIFB1	1.16	1.50	.24	2/24 hyp/error
DIFB2	5.13	5.19	.02*	2/24
DIFB3	8.02	5.72	.01*	2/24
DIFB4	.52	3.41	.05*	2/24
DIFB5	.21	.34	n.s.	2/24
DIFB6	4.21	1.28	.30	2/24
DIFB7	.41	.83	n.s.	2/24

* Significant
\# Approaching significance

Table 25

Discriminant Analysis: 02 Difference Power Test Variables

First Canonical Variate - 3 Groups - 82.5% of Canonical Variation

	Discriminant Function Coefficients for Variables							Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-.4136	-.4691	-.4530	-1.5928	.0058	-.0362	-.2259	.74	-1.19	-1.18
Standardized	-.3632	-.4665	-.5363	-.6199	.0046	-.0656	-.1542	--	--	--

Second Canonical Variate - 3 Groups - 17.5% of Canonical Variation

Raw	.5440	.5178	-.7122	.4762	-.2004	-.0051	-.3665	-.38	-.81	.12
Standardized	.4777	.5149	-.8432	.1853	-.1588	-.0093	-.2501	--	--	--

I vs. SA Canonical Variate

Raw	.5172	.4877	-.7390	.3766	-.1997	-.0073	-.3798	--	-.88	.04
Standardized	.4543	.4850	-.8748	.1465	-.1582	-.0113	-.2592	--	--	--

I vs. AS Canonical Variate

Raw	-.5386	-.5856	-.2565	-1.6616	.0567	-.0337	-.1250	.81	--	-1.12
Standardized	-.4730	-.5823	-.3036	-.6467	.0449	-.0611	-.0853	--	--	--

Table 25 (Continued)

AS vs. SA Canonical Variate

	Discriminate Function Coefficients for Variables							Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-.2814	-.3413	-.6010	-1.4458	-.0393	-.0364	-.3022	.63	-1.34	--
Standardized	-.2471	-.3394	-.7114	-.5627	-.0311	-.0660	-.2062	--	--	--

Bartlett's Chi Square Test of Significance

Variate(s)	χ^2	df	p
First and Second-3 Groups	21.22	14	.10#
Second-3 Groups	4.74	6	.58
I vs. AS	13.80	7	.06*
I vs. SA	4.64	7	.71
AS vs. SA	14.26	7	.05*

* Significant
\# Approaching significance

The magnitudes of the standardized coefficients of the first discriminate function indicated that the variables that contributed most to the differences, in order from most to least, are DIFB3, DIFB4, DIFB2, DIFB1 and DIFB7. The variables DIFB5 and DIFB6 contributed very little to the discrimination. The signs of the coefficients were all negative indicating that there was only one factor in the discrimination among the groups. The values of the first function for the groups AS, SA, AND I were .74, -1.19, and -1.18, respectively, indicating that with respect to performance for these seven power test variables groups SA and I were similar to each other and different from group AS.

The canonical variate for the pair of groups AS and SA was significant ($p < .05$), as was the variate for the groups I and AS ($p < .06$). These two variates are very similar to the first canonical variate of the overall groups differences. The discriminant function for the pair of groups I and SA was not significant.

Overall groups univariate analysis - 02 difference power test variables. The significance of the differences among the three groups for each of the 02 variables DIFB1 through DIFB7 was tested by analysis of variance; the statistics for the univariate F tests are reported in Table 24. The F tests for the 02 variables DIFB2, DIFB3, and DIFB4 were significant. These were the same variables that contributed most to the first discriminant function of the multivariate space.

Pairs of groups analysis - O2 difference power test variables. The statistics for the multivariate and univariate analyses of variance for the means of pairs of groups for the variables DIFB1 through DIFB7 are given in Table 26. The means of groups SA and I were not significantly different ($F = .65$), however, the multivariate differences between groups AS and I were ($p < .06$), as were the differences between groups AS and SA ($p < .05$). Univariate differences between groups AS and I were significant for the variables DIFB1, DIFB2, and DIFB4, and approached significance for DIFB3. Univariate differences between groups AS and SA were significant for the variables DIFB2 and DIFB3 and approached significance for the variables DIFB4 and DIFB6. While the multivariate differences between groups SA and I were not significant, the differences were significant for the variable DIFB3.

Summary of O2 Difference Power Test Variable Comparisons

The multivariate test of overall group differences for the O2 variables DIFB1 through DIFB7 approached significance ($p < .11$). Though the hypothesis H4.02 could not be rejected, the subsequent analyses were done as if it had been. The discriminant analyses, the analyses of pairs of groups, and the univariate analyses all indicate that if any differences existed, they were between group AS and the other two groups. Groups SA and I could not be distinguished from one another by these analyses. In general, the means of groups

Table 26

Pairs of Groups MANOVA and ANOVA for 02 Difference Power Test Variables

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	2.47	.06*	7/18 N/D
DIFB1	2.29	2.97	.10*	1/24 hyp/error
DIFB2	8.91	9.00	.01*	1/24
DIFB3	3.39	2.42	.14#	1/24
DIFB4	1.02	6.73	.02*	1/24
DIFB5	.28	.44	n.s.	1/24
DIFB6	2.16	.66	n.s.	1/24
DIFB7	.30	.65	n.s.	1/24

I vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	.65	n.s.	7/18 N/D
DIFB1	.32	.42	n.s.	1/24 hyp/error
DIFB2	.14	.14	n.s.	1/24
DIFB3	5.14	3.67	.07*	1/24
DIFB4	.14	.94	n.s.	1/24
DIFB5	.01	.01	n.s.	1/24
DIFB6	2.29	.70	n.s.	1/24
DIFB7	.14	.31	n.s.	1/24

Table 26 (Continued)

AS vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	2.59	.05*	7/18 N/D
DIFB1	.96	1.25	.28	1/24 hyp/error
DIFB2	6.91	6.98	.02*	1/24
DIFB3	15.83	11.30	.01*	1/24
DIFB4	.43	2.82	.11#	1/24
DIFB5	.38	.60	n.s.	1/24
DIFB6	8.38	2.55	.13#	1/24
DIFB7	.82	1.76	.20	1/24

* Significant
\# Approaching significance

was employed. Absences are considered to be random across groups. AS were less than those of the other two groups and negative for four of the seven variables. The differences involved the variables DIFB1, DIFB2, DIFB3, and DIFB4; these were the variables for which group AS had negative means.

Retention Test Data from Observation 03

The data from the observation 03 provided the basis for comparisons of group performances two weeks after the end of instruction. The 03 data do not provide measures of the amount of change that may have occurred during the two week period as the groups were known to be different at the beginning of the two week period. However, the 03 - 02 data that are presented in the section following this did provide direct measures of the amount of change that took place during the two week retention period.

The group means and standard errors for the 03 variables were estimated using the randomized block design and the scores of 13 subjects from group AS, 16 subjects from group SA, and 14 subjects from group I; the estimated means and standard errors are given in Table 27. These means and the analyses relating to them are discussed in the following order: the achievement variables first, then the total power test variables, and finally the difference power test variables.

The 03 Achievement Variables

The 03 achievement variables were measures of student performance under timed conditions two weeks after the termination of instruction. There were six basic achievement variables, +F, -F, ANC, C, SNB, and B and four constructed achievement variables, SUM, OP, RGP, and INT, that entered into the description of group performance following instruction.

The estimated group means for the 03 achievement variables are presented in Table 27. The null hypothesis to which these data apply is as follows.

H2.03 There are no differences among the group means for the achievement variables of 03.

This hypothesis was tested by multivariate analysis of covariance of the mean scores for the 03 variables +F, -F, ANC, C, SNB, and B using the covariates MET1 and 01SNB. This test is reported following the description of the group means.

Table 27

Estimated Group Means and Standard Errors for 03 Variables

Variable	Group (N)					
	AS (13)		SA (16)		I (14)	
	Estimated Mean	Estimated Standard Error	Estimated Mean	Estimated Standard Error	Estimated Mean	Estimated Standard Error
+F	8.49	.79	5.88	.66	4.51	.77
-F	6.67	.95	4.08	.79	5.06	.93
ANC	2.55	.47	2.35	.39	2.42	.46
C	3.93	.87	2.78	.72	4.05	.85
SNB	2.54	.43	1.22	.36	2.44	.42
B	2.60	.80	1.00	.66	2.36	.78
SUM	2.09	.37	1.36	.32	2.01	.35
OP	.34	.32	1.01	.27	.41	.39
RGP	.92	.30	.83	.25	.82	.28
INT	-.33	.12	.12	.11	-.43	.12
SUMB1	1.04	.45	1.20	.38	.71	.44
SUMB2	5.70	.77	4.61	.65	6.53	.76
SUMB3	4.21	.78	2.80	.65	3.94	.76
SUMB4	7.33	.24	7.47	.20	7.67	.24
SUMB5	5.76	.73	5.87	.61	6.76	.71
SUMB6	3.57	.84	3.01	.70	4.26	.82
SUMB7	5.95	.67	6.05	.56	6.25	.65
DIFB1	-.13	.14	.21	.12	-.20	.14
DIFB2	.32	.36	.88	.30	.63	.35
DIFB3	.75	.48	.94	.40	1.22	.47
DIFB4	.15	.16	-.08	.13	-.36	.15
DIFB5	.85	.28	.70	.24	.43	.28
DIFB6	.83	.51	1.35	.43	.56	.50
DIFB7	.37	.27	.52	.22	.33	.26

The group means for the achievement variables of 03 adjusted by the covariates MET1 and 01SNB are listed in Table 28. Though the overall effect of the adjustments was to lower the estimated means of group I with respect to the other two groups, the adjustments were not large. The order of the group means was changed for only the variables ANC and C.

TABLE 28

Adjusted Group Means for the 03 Achievement Variables
Covariates: MET1, 01SNB

Groups(N)	Variables									
	+F	-F	ANC	C	SNB	B	SUM	OP	RGP	INT
AS(13)	8.63	6.60	2.71	4.32	2.65	2.81	2.23	.40	.90	-.35
SA(16)	6.21	4.47	2.51	3.07	1.40	1.30	1.52	1.00	.87	.11
I(14)	4.04	4.72	2.09	3.37	2.14	1.85	1.71	.35	.81	-.40

The adjusted means of group AS exceeded those of the other two groups for all six of the basic achievement variables of 03. The adjusted means of group SA were the least for the variables -F, C, SNB, and B, while the means of group I were less than those of the other two groups for the variables +F and ANC.

The adjusted group means for the constructed achievement variables SUM, OP, RGP, and INT organized the information from the variables ANC, C, SNB, and B in the following way. The total

performance as measured by the variable SUM was greatest for group AS and least for group SA. Though the adjusted means for the variable OP were all positive, group AS had the least mean and group SA the greatest. The adjusted group means for the variable RGP were similar and positive for all three groups.

The adjusted means of groups AS and I were negative for the variable INT and of about the same magnitude, while the mean of group SA was positive. A positive mean for the variable INT indicates that the difference between SNB and B is less than the difference between ANC and C.

The statistics for the multivariate F test of hypothesis H2.03 are given in Table 29. The overall F was significant ($p < .002$) and the hypothesis accordingly rejected. Therefore, the analyses outlined in Figure 8B were undertaken to specify the differences.

Discriminant analysis for the 03 achievement variables. The results of the discriminant analyses of the group means for the 03 achievement variables +F, -F, ANC, C, SNB, and B are given in Table 30. The first canonical variate for the means of all three groups accounted for 72.1% of the canonical variation, and the second the remaining 17.9%. Since both contrasts taken together provided a significant contrast among the groups ($p < .001$), and the second contrast taken alone was also significant ($p < .05$), both of the canonical variates are considered further.

The magnitudes of the standardized coefficients of the first contrast indicated the variables that contributed most to the differences listed

TABLE 29

Overall Groups MANCOVA and ANCOVAs for 03 Achievement Variables
Covariates: MFT 1, 01SNB

Variables	Hypothesis Mean Squares	F	p	df	
+F, -F, ANC, C, SNB, B	(multivariate)	3.83	.002*	12/32	N/D
+F	57.31	16.06	.001*	2/21	hyp/error
-F	15.42	2.92	.08*	2/21	
ANC	1.10	.80	n.s.	2/21	
C	4.96	1.17	.33	2/21	
SNB	4.90	5.44	.02*	2/21	
B	7.03	1.69	.21	2/21	
SUM	1.57	3.14	.07*	2/21	
OP	1.67	1.44	.26	2/21	
RGP	.02	.02	n.s.	2/21	
INT	1.04	5.86	.01*	2/21	

* Significant

in order from most to least, were +F, -F, and SNB. The variables ANC, C, and B contributed very little to this contrast. The coefficients for the variables +F and SNB were positive, while the coefficient for the variable -F was negative. Hence, differences among the groups were greatest when subtraction facts performance was contrasted with performance for both addition facts and subtraction without regrouping. The values of the

first function for the groups AS, SA, and I were 5.47, 3.74, and 2.58, respectively. While a simple descriptive label for this contrast does not seem to be available, it is clear that this contrast involves the variables +F, -F, and SNB, and that it was primarily a contrast between groups AS and I; group SA appears to be more similar to group I than it does to group AS, however.

The magnitudes of the standardized coefficients for the second contrast indicated variables that contributed most to the differences, listed in order from most to least, were SNB, ANC, C, and B; the variables +F and -F were relatively unimportant for this contrast. The coefficients for SNB and C were negative and the coefficients for ANC and B were positive; this contrast can therefore be identified as a contrast based upon an interaction of operation with regrouping. This contrast is similar to that provided by the constructed variable INT. The values of the second discriminant function for the groups AS, SA, and I were -1.69, -.36, and -1.89, respectively; these values indicate that the second contrast was a contrast of group SA with both of the other groups.

The discriminant analyses for the pairs of groups help to clarify the situation somewhat. The canonical variates for the pairs of groups I and AS, and AS and SA were similar to the first canonical variate for overall groups (see Table 30), indicating that the first canonical variate was indeed a contrast of group AS with the other two groups. While all six of the variables were important in the discriminant function for groups I and SA, the signs of the coefficients

Table 30

Discriminant Analysis: 03 Achievement Variables Adjusted Means

Covariates: MET1, O1SNB

First Canonical Variate - 3 Groups - 72.1% of Canonical Variation

	Discriminant Function Coefficients for Variables						Function Values for Groups		
	+F	-F	ANC	C	SNB	B	AS	SA	I
Raw	.7039	-.2691	-.0957	.0378	.5990	-.1122	5.47	3.74	2.58
Standardized	1.3295	-.6186	-.1122	.0776	.5686	-.2288	--	--	--

Second Canonical Variate - 3 Groups - 17.9% of Canonical Variation

Raw	.1540	-.1662	.7778	-.4176	-1.2126	.3513	-1.69	-.36	-1.89
Standardized	.2910	-.3820	.9120	-.8581	-1.1510	.7168	--	--	--

I vs. AS Canonical Variate

Raw	-.7127	.2797	.0429	-.0094	-.5157	.0882	-5.35	--	-2.44
Standardized	-1.3462	.6430	.0504	-.0194	-.4895	.1798	--	--	--

I vs. SA Canonical Variate

Raw	.5488	-.2952	.5610	-.3095	-.6023	.2117	--	1.98	.06
Standardized	1.0366	-.6786	.6579	-.6360	-.5718	.4319	--	--	--

Table 30 (Continued)

AS vs. SA Canonical Variate

Discriminant Function Coefficients for Variables							Function Values for Groups		
	+F	-F	ANC	C	SNB	B	AS	SA	I
Raw	-.4650	.1125	.5490	-.2840	-1.2130	.3027	-5.37	-3.18	--
Standardized	-.8784	.2587	.6438	-.5835	-1.1514	.6176	--	--	--

Bartlett's Chi Square Test of Significance

Variate(s)	χ^2	df	P
First and Second-3 Groups	32.95	12	.001*
Second-3 Groups	11.40	5	.05*
I vs. AS	20.88	6	.002*
I vs. SA	13.35	6	.04*
AS vs. SA	15.67	6	.02*

* Significant

indicate that this contrast is a mixture of the interaction noted for the second canonical variate for overall groups and the addition facts versus subtraction facts contrast of the first canonical variate.

Overall groups univariate analysis - 03 achievement variables. The next step in the description of group differences for the 03 achievement variables was the testing of differences among the three groups for each individual variable. The statistics for the F tests of these differences are given in Table 29. Significant differences existed among the groups for the variables +F ($p < .001$), -F ($p < .08$), SNB ($p < .02$), SUM ($p < .07$), and INT ($p < .01$).

Pairs of groups analysis - 03 achievement variables. The multivariate and univariate differences between pairs of groups for the 03 achievement variables were tested for significance; the statistics relating to these tests are reported in Table 31. The multivariate F was significant for the pairs of groups I and AS ($p < .003$), I and SA ($p < .04$), and AS and SA ($p < .02$); hence, the three groups were different from one another. The univariate differences between groups I and AS were significant for the variables +F ($p < .001$) and - F ($p < .08$), and approached significance for the variable SUM ($p < .11$). The univariate differences between groups I and SA were significant for the variables +F ($p < .01$), SNB ($p < .07$), and INT ($p < .01$), and approached significance for the variable OP ($p < .15$). Finally, the univariate differences between groups AS and SA were significant for +F ($p < .01$), -F ($p < .04$), SNB ($p < .01$), B ($p < .09$), SUM ($p < .03$), and INT ($p < .02$), and approached significance for C ($p < .15$).

Table 31

MANCOVAs and ANCOVAs for 03 Achievement Variables by Pairs of Groups

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	5.84	.003*	6/16 N/D
+F	114.57	32.11	.001*	1/21 hyp/error
-F	19.15	3.62	.08*	1/21
ANC	2.07	1.51	.24	1/21
C	4.91	1.16	.30	1/21
SNB	1.42	1.57	.23	1/21
B	4.99	1.20	.29	1/21
SUM	1.47	2.95	.11#	1/21
OP	.01	.01	n.s.	1/21
RGP	.04	.04	n.s.	1/21
INT	.01	.10	n.s.	1/21

I vs. SA

+F, -F, ANC, C, SNB, B	(multivariate)	2.93	.04*	6/16 N/D
+F	24.43	8.25	.01*	1/21 hyp/error
-F	.39	.07	n.s.	1/21
ANC	1.12	.82	n.s.	1/21
C	.56	.13	n.s.	1/21
SNB	3.41	3.78	.07*	1/21
B	1.96	.47	n.s.	1/21
SUM	.22	.43	n.s.	1/21
OP	2.61	2.25	.15#	1/21
RGP	.02	.02	n.s.	1/21
INT	1.66	9.41	.01*	1/21

Table 31 (Continued)

AS vs. SA

+F, -F, ANC, C, SNB, B	(multivariate)	3.70	.02*	6/16 N/D
+F	35.66	10.00	.01*	1/21 hyp/error
-F	27.52	5.21	.04*	1/21
ANC	.23	.17	n.s.	1/21
C	9.49	2.25	.15#	1/21
SNB	9.47	10.51	.	1/21
B	14.00	3.36	.09*	1/21
SUM	3.03	6.07	.03*	1/21
OP	2.15	1.85	.19	1/21
RGP	.01	.01	n.s.	1/21
INT	1.29	7.28	.02*	1/21

* Significant
\# Approaching significance

Summary of 03 Achievement Variable Comparisons

The multivariate differences among the three groups for the 03 achievement variables +F, -F, ANC, C, ANB, and B were significant, and hence the hypothesis H2.03 is rejected. When the groups were considered by pairs, all of the multivariate differences were significant. Hence, each of the three groups was different from each of the others.

The adjusted means of group AS were greater than those of either of the other two groups for the six basic achievement variables, though the patterns of differences were not the same. The first canonical variate for the means of all three groups was primarily a contrast of group AS with both groups SA and I, while the second was a

contrast of group SA with groups AS and I. The important variables in the first contrast were +F, -F, and SNB, while the important variables in the second contrast were ANC, C, SNB, and B.

The means of groups AS and SA differed significantly for the individual variables +F, -F, SNB, B, SUM, and INT, while those of groups AS and I differed significantly for the individual variables +F, -F, SNB, B, SUM, and INT, while those of groups AS and I differed significantly only for +F and -F. However, the difference between groups AS and I for the variable SUM approached significance as well.

The means of groups I and SA differed significantly for the variables SNB, +F, and INT, while the group differences for the variable SUM approached significance. The means of group I were greater than those of group SA for SNB and SUM, but were less than those of group SA for +F and INT.

The significant differences between group SA and each of the other two groups for the variable INT were reflected in the second overall groups discriminant function. This is correlated with a generally low performance of group SA for both of the subtraction variables SNB and B.

The 03 Total Power Test Variables

The 03 total power test variables SUMB1 through SUMB7 were measures of total power test performance two weeks after instruction was terminated. The estimated group means for these variables are reported in Table 27. The null hypothesis to which these data apply is as follows.

H3.03 There are no differences among the group means for the power test measures of total performance at 03.

This hypothesis was tested by multivariate analysis of covariance of the mean scores for the 03 variables SUMB1 through SUMB7 using covariate MET2. This test is reported following the description of the group means.

The adjusted group means for the 03 total power test variables are listed in Table 32. When these means are compared to the estimated group means of Table 27, the overall changes in the magnitudes of the means do not appear to be large; in no case was the order of the group means altered by the addition of the covariate to the model.

The overall differences among the means for the groups appear to be slight though a trend is apparent in that the means of group I for the variables SUMB2 and SUMB4 through SUMB7 exceeded those of the other two groups.

The statistics for the multivariate F test of the hypothesis H3.03 are reported in Table 33. The multivariate F ratio was .58 and hence the hypothesis is not rejected. The subsequent multivariate and univariate tests of group differences are reported in Tables 33 and 34. Overall group differences for the variable 03SUMB2 approached significance ($p < .15$) and the differences between groups I and SA for this variable were significant ($p < .05$). However, all other tests failed to indicate any group differences for the 03 total power test variables.

Table 32

Adjusted Group Means for the 03 Total Power Test Variables
Covariate: MET2

Groups(N)	Variables						
	SUMB1	SUMB2	SUMB3	SUMB4	SUMB5	SUMB6	SUMB7
AS(13)	.99	5.63	4.14	7.32	5.73	3.50	5.94
SA(16)	1.25	4.69	2.88	7.48	5.91	3.09	6.06
I(14)	.70	6.52	3.93	7.67	6.76	4.25	6.24

Table 33

Overall Groups MANCOVA and ANCOVAs for 03 Total Power Test Variables
Covariates: MET 2

Variables	Hypothesis Mean Squares	F	p	df	
SUMB1, SUMB2, SUMB3, SUMB3, SUMB5, SUMB6, SUMB7	(multivariate)	.58	n.s.	14/32	N/D
SUMB1	1.00	.64	n.s.	2/22	hyp/error
SUMB2	11.17	2.15	.15#	2/22	
SUMB3	6.01	1.11	.35	2/22	
SUMB4	.36	.55	n.s.	2/22	
SUMB5	3.71	.64	n.s.	2/22	
SUMB6	4.52	.72	n.s.	2/22	
SUMB7	.28	.05	n.s.	2/22	

Approaching significance

Table 34

Pairs of Groups MANCOVAs and ANCOVAs for 03 Total Power Test Variables

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.54	n.s.	7/16 N/D
SUMB1	.48	.31	n.s.	1/22 hyp/error
SUMB2	4.62	.89	n.s.	1/22
SUMB3	.26	.05	n.s.	1/22
SUMB4	.71	1.10	.31	1/22
SUMB5	6.23	1.07	.32	1/22
SUMB6	3.26	.52	n.s.	1/22
SUMB7	.55	.11	n.s.	1/22

I vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.70	n.s.	7/16 N/D
SUMB1	2.01	1.29	.26	1/22 hyp/error
SUMB2	22.29	4.29	.05*	1/22
SUMB3	7.39	1.36	.26	1/22
SUMB4	.23	.35	n.s.	1/22
SUMB5	4.80	.82	n.s.	1/22
SUMB6	8.90	1.43	.25	1/22
SUMB7	.22	.04	n.s.	1/22

Table 34 (Continued)

AS vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	.49	n.s.	7/16 N/D
SUMB1	.43	.28	n.s.	1/22 hyp/error
SUMB2	5.58	1.07	.31	1/22
SUMB3	9.96	1.84	.19	1/22
SUMB4	.17	.26	n.s.	1/22
SUMB5	.20	.03	n.s.	1/22
SUMB6	1.07	.17	n.s.	1/22
SUMB7	.10	.02	n.s.	1/22

* Significant

The 03 Difference Power Test Variables

The 03 difference power test variables DIFB1 through DIFB7 were power test measures of differences in performance for similar addition and subtraction behaviors two weeks after instruction was terminated. It should be recalled that a negative score for one of these variables indicates that for a pair of similar behaviors the subtraction score was greater than the addition score.

The estimated group means for the 03 difference power test variables are found in Table 27. The null hypothesis to which these data apply is as follows.

H4.03 There are no differences among the group means for the 03 power test measures of differential performance for addition and subtraction.

Since no covariates were entered into the analysis model for estimating the group means for the variables DIFB1 through DIFB7, hypothesis H4.03 was tested by multivariate analysis of variance of the estimated group means for the 03 variables DIFB1 through DIFB7. This test is reported following the description of the group means.

Of the 21 means for the variables DIFB1 through DIFB7, four were negative. Two of these four were group I means, one was a group AS mean and one was a group SA mean. The preponderance of positive scores indicates that addition performance was generally superior to subtraction performance at 03. The means of group I were less than those of the other groups for five of the seven variables, and for two of these five variables the means of group I were negative. All of this suggests that the generally better performance for addition was least pronounced for group I.

The statistics for the F test of the hypothesis H4.03 are reported in Table 35. The multivariate F for hypothesis H4.03 is significant ($p < .10$) and the hypothesis is therefore rejected. Therefore, the analyses outlined in Figure 8B were undertaken to specify the differences.

Discriminant analysis for 03 difference power test variables. The results of the discriminant analyses of the group means for the 03 variables DIFB1 through DIFB7 are given in Table 36. The first canonical variate for the means of the three groups accounted for 75.1% of the canonical variation, and the second the remaining 24.9%.

Table 35

Overall Groups MANOVA and ANOVAs for 03 Difference Power Test Variables

Variables	Hypothesis Mean Squares	F	p	df	
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	1.74	.10*	14/34	N/D
DIFB1	.64	2.93	.08*	2/23	hyp/error
DIFB2	.98	.69	n.s.	2/23	
DIFB3	.65	.25	n.s.	2/23	
DIFB4	.77	2.81	.09*	2/23	
DIFB5	.55	.61	n.s.	2/23	
DIFB6	.14	.74	n.s.	2/23	
DIFB7	.13	.16	n.s.	2/23	

* Significant

While both of the contrasts taken together provided a significant contrast among the groups ($p < .09$), the second contrast by itself did not ($p < .38$). The first canonical variate is therefore taken as the best single descriptor of the multivariate group differences.

The magnitudes of the standardized coefficients of the first discriminant function indicated the variables contributing most to the differences, listed in order from most to least, were DIFB3, DIFB4, DIFB2, DIFB6, DIFB1, and DIFB5. The variable DIFB7 was not important to this contrast. The coefficients for the variables

Table 36

Discriminant Analysis: 03 Difference Power Test Variables

First Canonical Variate-Groups - 75.1% of Canonical Variation

	Discriminant Function Coefficients for Variables							Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-1.093	.6346	.7705	-2.1047	-.4662	-.3474	-.0113	-.08	.42	1.92
Standardized	-.5119	.7586	1.2396	-1.1012	-.4426	-.5924	-.0102	--	--	--

Second Canonical Variate-3 Groups - 24.9% of Canonical Variation

Raw	-1.7300	-.4564	.1640	.7552	.2899	-.2709	.1901	.41	-.56	.02
Standardized	-.8104	-.5456	.2638	.3956	.2753	-.4619	.1708	--	--	--

I vs. AS Canonical Variate

Raw	-.7418	.7102	.7249	-2.2104	-.5130	-.2892	-.0475	-.16	--	1.88
Standardized	-.3475	.8489	1.1663	-1.1565	-.4871	-.4931	-.0427	--	--	--

I vs. SA Canonical Variate

Raw	-1.7560	.3602	.7615	-1.5366	-.2851	-.4322	..0757	--	.04	1.72
Standardized	-.8226	.4306	1.2251	-.8040	-.2707	-.7371	.0680	--	--	--

Table 36 (Continued)

AS vs. SA Canonical Variate								Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-1.1412	-.6739	-.1613	1.5412	.4535	-.1075	.1785	..41	-.84	--
Standardized	-.5346	-.8055	-.2594	.8068	.4306	-.1832	.1604	--	--	--

Bartlett's Chi Square Test of Significance

Variate(s)	χ^2	df	p
First and Second-3 Groups	21.58	14	.09*
Second-3 Groups	6.40	6	.38
I vs. AS	14.07	7	.05*
I vs. SA	11.52	7	.12#
AS vs. SA	6.89	7	.44

* Significant
\# Approaching significance

DIFB2 and DIFB3 were positive, forming one pole of the contrast, while the coefficients for the other variables were negative and formed the other pole of the contrast. While DIFB2 and DIFB3 were concerned with expanded notation algorithms and so identify that pole, the identification of the other pole is not clear. The values of this function for the groups AS, SA, and I were -.08, .42, and 1.92, respectively; hence, the contrast was primarily between groups I and AS with group SA more closely resembling group AS than group I.

The discriminant analyses for each pair of groups also indicated a general similarity of groups AS and SA. The contrast for the pair of groups AS and SA was not significant, while the contrast for the pair I and AS was ($p < .05$), and the contrast for the pair I and SA approached significance ($p < .12$). The canonical variates for the pairs of groups I and AS, and I and SA appear to be similar to the first canonical variate for all groups.

Overall groups univariate analysis - 03 difference power test variables. The significance of the differences among the three groups for each of the 03 variables DIFB1 through DIFB7 was tested by analysis of variance; the statistics for these tests are reported in Table 35. The F tests were found to be significant for the variables DIFB1 ($p < .08$) and DIFB4 ($p < .09$) only.

Pairs of groups analysis - 03 difference power test variables. Finally, the multivariate and univariate differences between pairs of groups for the 03 difference power test variables were also

tested for significance. The statistics relating to these tests are reported in Table 37. The multivariate differences between groups I and AS were significant ($p < .06$), with significant univariate differences for the variable DIFB4 ($p < .03$) only. The multivariate differences between groups I and SA approached significance ($p < .13$) with significant univariate differences for the variable DIFB1 ($p < .04$), and the multivariate differences between groups AS and SA were not significant. However, the difference between group AS and SA was significant for the single variable DIFB1 ($p < .09$).

Summary of 03 Difference Power Test Variable Comparisons

The multivariate differences among the three groups for the 03 difference power test variables were found to be significant ($p < .10$) and the hypothesis H4.03 was rejected. The discriminant analysis indicated that the difference was primarily between groups I and AS with group SA most closely resembling group AS. The most important variable in the contrast between these groups is DIFB4, however, the general trend was for the means of group I to be less than those of the other two groups. It should be noted that none of the multivariate tests found to be significant at the .10 level were significant below the .05 level.

03 - 02 Data

Measures of changes in performance during the retention period were provided by the difference in means from 02 to 03. While the

Table 37

Pairs of Groups MANCOVAs and ANCOVAs for 03 Difference Power Test Variables

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	2.57	.06*	7/17 N/D
DIFB1	.03	.13	n.s.	1/23 hyp/error
DIFB2	.58	.41	n.s.	1/23
DIFB3	1.28	.49	n.s.	1/23
DIFB4	1.52	5.57	.03*	1/23
DIFB5	1.05	1.16	.29	1/23
DIFB6	.43	.15	n.s.	1/23
DIFB7	.01	.01	n.s.	1/23

I vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	1.96	.13#	7/17 N/D
DIFB1	1.11	5.07	.04*	1/23 hyp/error
DIFB2	.41	.28	n.s.	1/23
DIFB3	.51	.20	n.s.	1/23
DIFB4	.51	1.87	.19	1/23
DIFB5	.50	.55	n.s.	1/23
DIFB6	4.11	1.41	.25	1/23
DIFB7	.23	.29	n.s.	1/23

Table 37 (Continued)

AS vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7,	(multivariate)	1.03	.45	7/17 N/D
DIFB1	.72	3.29	.09*	1/23 hyp/error
DIFB2	1.97	1.38	.26	1/23
DIFB3	.22	.09	n.s.	1/23
DIFB4	.33	1.22	.29	1/23
DIFB5	.14	.15	n.s.	1/23
DIFB6	1.67	.57	n.s.	1/23
DIFB7	.15	.18	n.s.	1/23

* Significant
\# Approaching significance

primary use of the 02 and 03 data was to assess group performance following instruction, and assessment of forgetting was provided by using the 02 and 03 data in concert. Since the resulting 03 - 02 data are not independent from the 02 and 03 data, this analysis is regarded as supplementary to the 02 and 03 data analyses. This section presents an analysis of the 03 - 02 difference scores for the three groups.

The estimated group means for the 03 - 02 variables were obtained by using the randomized block model and the scores of 12 subjects from group AS, 15 subjects from group SA and 14 subjects from group I that had complete data for both observation 02 and 03. The estimated means

are reported in Table 38. These means and the analyses relating to them are discussed the following order; the achievement variables first, then the total power test variables, and finally the difference power test variables.

TABLE 38

Estimated Group Means for 03-02 Variables

Achievement Variables

Groups(%)	+F	-F	ANC	C	SNB	B	SUM	OP	R6P	INT
AS(12)	.39	.26	.43	.22	.17	-.54	.10	.34	.40	-.09
SA(15)	1.12	.38	.22	.16	-.55	.39	-.02	.33	-.30	.44
I(14)	-.35	1.26	-.08	-.05	.04	-.32	-.06	.00	.07	-.13

Sum Power Test Variables

	SUMB1	SUMB2	SUMB3	SUMB4	SUMB5	SUMB6	SUMB7
AS(12)	-.79	.66	1.17	.14	-.05	-.02	-.05
SA(15)	-.23	-.54	-.42	-.17	-.67	-.11	-.42
I(14)	-.12	[illegible]	.06	.94	.52	-.66	.44

Difference Power Test Variables

	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7
AS(12)	.47	.23	.81	.33	.55	1.04	.05
SA(15)	.32	.17	-.36	-.12	.09	.33	.04
I(14)	-.36	-.41	.80	-.51	.03	.01	-.08

The 03 - 02 Achievement Variables

After a period of no instruction or practice, performance was expected to be lower on each of the achievement variables. However, with respect to the six basic achievement variables, the means of group I increased for the variables -F and SNB, the means of group AS increased for all of the variables except B, and the means of group SA increased for all of the variables but SNB. Since the interest was in possible differences in the rates of forgetting between the groups, there were no tests made of the null hypothesis that the group means were zero.

The null hypothesis to which the 03 - 02 achievement variables apply is as follows.

H2.03 - 02 There are no group differences in the rate of forgetting for the achievement variables.

This hypothesis was tested by multivariate analysis of covariance on the mean scores for the 03 - 02 variables +F, -F, ANC, C, SNB, and B using the covariates MET1 and 01SNB. The group means for the achievement variables of 03 - 02 adjusted by the covariates MET1 and 01SNB are listed in Table 39; the adjustments appear to be minor.

Table 40 reports the statistics of the multivariate F test for the hypothesis H2.03 - 02. The value of F was not significant ($p < .22$), and the hypothesis is not rejected. Table 40 also reports the statistics for the univariate tests of overall group differences. While none of the F ratios were significant, they approached significance for the variables

TABLE 39

Adjusted Group Means for 03-02 Achievement Variables
Covariates: MET1, 01SNB

Groups(N)	Variables									
	+F	-F	ANC	C	SNB	B	SUM	OP	RGP	INT
AS(12)	.26	.28	.51	.43	.16	-.53	.15	.42	.36	-.07
SA(15)	1.00	.25	.22	.12	-.51	.28	-.02	.33	-.24	.41
1(14)	-.10	1.37	-.18	-.22	.01	-.32	-.11	-.07	.05	-.12

TABLE 40

Overall Groups MANCOVA and ANCOVAs for 03-02 Achievement Variables
Covariates: MET1, 01SNB

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	1.42	.22	12/28 N/D
+F	3.77	1.47	.25	2/19 hyp/error
-F	4.40	1.10	.35	2/19
ANC	1.21	.87	n.s.	2/19
C	1.06	.69	n.s.	2/19
SNB	1.45	2.33	.13#	2/19
B	2.07	.76	n.s.	2/19
SUM	.18	.63	n.s.	2/19
OP	.70	1.04	.37	2/19
R6P	1.02	1.29	.30	2/19
INT	1.00	2.22	.14#	2/19

Approaching significance

SNB ($p < .13$) and INT ($p < .14$). These variables are of particular interest since overall group differences for these two variables were found at 03, but not at 02. The multivariate and univariate differences between pairs of groups for the 03 - 02 achievement variables were also tested for significance. The statistics relating to these tests are reported in Table 41. Significant multivariate differences were found to exist between the pair of groups I and SA ($p < .08$) only. The univariate differences between these two groups were significant for the variable INT ($p < .08$) and approached significance for the variables +F ($p < .12$) and SNB ($p < .13$). While the multivariate differences between groups AS and SA were not significant, significant differences were indicated for the single variable SNB ($p < .06$), and differences approached significance for the variables RGP ($p < .13$) and INT ($p < .12$). The were no significant differences between groups I and AS.

Summary: 03 - 02 Achievement Test Variables

Though overall group differences in the rates of forgetting for the achievement variables cannot be inferred, there was some evidence of differences in the rate of forgetting between pairs of groups. While no differences were indicated between groups I and AS, the rates of forgetting for groups I and SA may be different for the variables +F, SNB, and INT, and there may also be differences between groups AS and SA in the rate of forgetting for the variables SNB, RGP, and INT. All of these possible differences are consistent with the changes in the pattern of group differences from 02 to 03.

Table 41

MANCOVAs and ANCOVAs for 03-02 Achievement Variables by Pairs of Groups

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	.78	n.s.	6/14 N/D
+F	.65	.25	n.s.	1/19 hyp/error
-F	5.96	1.49	.24	1/19
ANC	2.38	1.72	.21	1/19
C	2.11	1.37	.26	1/19
SNB	.11	.18	n.s.	1/19
B	.23	.08	n.s.	1/19
SUM	.35	1.21	.29	1/19
OP	1.19	1.77	.20	1/19
RGP	.48	.60	n.s.	1/19
INT	.02	.03	n.s.	1/19

I vs. SA

Variables	Hypothesis Mean Squares	F	p	df
+F, -F, ANC, C, SNB, B	(multivariate)	2.50	.08*	6/14 N/D
+F	7.10	2.77	.12#	1/19 hyp/error
-F	7.32	1.82	.20	1/19
ANC	.93	.67	n.s.	1/19
C	.67	.44	n.s.	1/19
SNB	1.59	2.56	.13#	1/19
B	2.10	.77	n.s.	1/19
SUM	.04	.15	n.s.	1/19
OP	.91	1.36	.26	1/19
RGP	.51	.64	n.s.	1/19
INT	1.61	3.58	.08*	1/19

Table 41 (Continued)

AS vs. SA

+F, -F, ANC, C, SNB, B	(multivariate)	1.29	.33	6/14 N/D
+F	3.10	1.21	.29	1/19 hyp/error
-F	.01	.00	n.s.	1/19
ANC	.46	.33	n.s.	1/19
C	.52	.34	n.s.	1/19
SNB	2.50	4.03	.06*	1/19
B	3.68	1.36	.26	1/19
SUM	.17	.56	.45	1/19
OP	.04	.06	.80	1/19
RGP	2.04	2.56	.13#	1/19
INT	1.24	2.75	.12#	1/19

* Significant
\# Approaching significance

The 03 - 02 Total Power Test Variables

The 03 - 02 total power test variables SUMB1 through SUMB7 are measures of the change in these variables from observation 02 to 03. The estimated group means for these variables are reported in Table 38. The null hypothesis to which these data apply is as follows.

H3.03 - 02 There are no group differences in the rate of forgetting for the power test measures of total performance.

This hypothesis is tested by multivariate analysis of covariance of the mean scores for the 03 - 02 variables SUMB1 through SUMB7 using the covariate MET2.

The adjusted group means for the 03 - 02 total power test variables are given in Table 42. There are no marked differences between these means and the estimated means presented in Table 38. The means of group SA are negative for all of these variables, while only three of the means of group AS are negative, and only two of the means of group I are negative. With some exceptions the order of the groups means from most to least is SA, AS, and I.

The statistics for the multivariate F test of hypothesis H3.03 - 02 are given in Table 43. The overall F was not significant ($p < .20$) and hence the hypothesis is not rejected. Further tests of the overall groups univariate differences are also reported in Table 43; significant differences were indicated for the variable SUMB2 only ($p < .09$).

The statistics relating to the tests of the differences between pairs of groups are reported in Table 44. While the multivariate test of differences between the groups I and SA approached significance ($p < .13$), the multivariate tests for the other pairs of groups were not significant. The univariate differences between groups I and SA were significant for the variables SUMB2 ($p < .04$), SUMB4 ($p < .09$), and SUMB5 ($p < .07$), and approached significance for the variable (SUMB7 ($p < .13$). The differences between groups AS and SA were significant for the variable SUMB3 ($p < .07$) only.

TABLE 42

Adjusted Group Means for 03-02 Total Power Test Variables
Covariate: MET2

Groups (N)	Variables						
	SUMB1	SUMB2	SUMB3	SUMB4	SUMB5	SUMB6	SUMB7
AS(12)	-.76	.62	1.15	.17	-.05	-.01	.01
SA(15)	-.26	-.51	-.40	-.20	-.66	-.12	-.47
I(14)	-.13	1.24	.06	.94	.52	-.66	.43

TABLE 43

Overall Groups MANCOVA and ANCOVAs for 03-02 Total Power Test Variables
Covariate: MET2

Variables	Hypothesis Mean Squares	F	p	df
SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	1.45	.20	14/28 N/D
SUMB1	1.17	1.05	.37	2/20 hyp/error
SUMB2	9.69	2.79	.09*	2/20
SUMB3	7.03	2.01	.16	2/20
SUMB4	4.05	1.67	.22	2/20
SUMB5	4.26	1.93	.18	2/20
SUMB6	1.36	.56	n.s.	2/20
SUMB7	2.50	1.26	.31	2/20

* Significant

Table 44

MANCOVAs and ANCOVAs for 03 - 02 Total Power Test Variables by Pairs of Groups

Covariate: MET2

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	1.36	.30	7/14 N/D
SUMB1	2.12	1.90	.19	1/20 hyp/error
SUMB2	2.08	.60	n.s.	1/20
SUMB3	6.44	1.84	.19	1/20
SUMB4	3.17	1.31	.27	1/20
SUMB5	1.72	.78	n.s.	1/20
SUMB6	2.26	.93	n.s.	1/20
SUMB7	.97	.49	n.s.	1/20

I vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	2.04	.13#	7/14 N/D
SUMB1	.10	.09	n.s.	1/20 hyp/error
SUMB2	18.72	5.40	.04*	1/20
SUMB3	1.29	.37	n.s.	1/20
SUMB4	7.91	3.26	.09*	1/20
SUMB5	8.49	3.84	.07*	1/20
SUMB6	1.77	.73	n.s.	1/20
SUMB7	4.98	2.51	.13#	1/20

Table 44 (Continued)

AS vs. SA

SUMB1, SUMB2, SUMB3, SUMB4, SUMB5, SUMB6, SUMB7	(multivariate)	1.05	.44	7/14 H/D
SUMB1	1.41	1.26	.28	1/20 hyp/error
SUMB2	7.19	2.07	.17	1/20
SUMB3	13.70	3.93	.07#	1/20
SUMB4	.77	.22	n.s.	1/20
SUMB5	2.12	.96	n.s.	1/20
SUMB6	.07	.03	n.s.	1/20
SUMB7	1.30	.65	n.s.	1/20

* Significant
# Approaching significance

Summary: 03 - 02 Total Power Test Variables

In summary, the differences among the means for the total power test variables of 03 - 02 indicated that group SA forgot the most and group I the least during the two-week retention period, however, the overall group differences in the rates of forgetting were not significant. There was some evidence of significant differences between groups SA and I, but the differences between groups AS and SA and between groups AS and I were not significant.

The 03 - 02 Difference Power Test Variables

The 03 - 02 difference power test variables DIFB1 through DIFB7 were power test measures of the change in the relationship between the addition power test measures and the subtraction power

test measures at 02 and 03. The estimated group means for these variables are found in Table 38. The null hypothesis to which these data apply is as follows.

H4.03 - 02 There are no group differences in the rates of change in the power test measures of operation differences.

This hypothesis is tested by multivariate analysis of variance of the estimated group means for the 03 - 02 variables DIFB1 through DIFB7.

The means of group AS for these variables are all positive and, generally, of greater magnitude than the means of the other two groups. These means reflect the fact that at observation 02 (see Table 16), the means of group AS for the difference power test variables were less than those of the other two groups, while at observation 03 (see Table 27), the means of all three groups appear more similar.

The statistics relating to the test of hypothesis H4.03 - 02 are reported in Table 45. The multivariate F for this test was significant ($p < .06$) and so the hypothesis is rejected. Therefore, the subsequent analysis as outlined in Figure 8B was carried out.

Discriminant analysis for 03 - 02 difference power test variables. The results of the discriminant analyses of the group means for the 03 - 02 variables DIFB1 through DIFB7 are presented in Table 46. The first canonical variate for the means of all three groups accounted for 73.7% of the canonical variation, and the second the remaining

TABLE 45

Overall Groups MANOVA and ANOVAs for 03-02 Difference Power Test Variables

Variables	Hypothesis Mean Squares	F	p	df
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	1.94	.06	14/30 N/D
DIFB1	2.16	1.71	.21	2/21 hyp/error
DIFB2	1.43	.91	n.s.	2/21
DIFB3	5.53	3.37	.06*	2/21
DIFB4	1.89	5.89	.01*	2/21
DIFB5	.87	.52	n.s.	2/21
DIFB6	2.97	2.28	.13#	2/21
DIFB7	.07	.13	n.s.	2/21

* Significant

\# Approaching significance

26.3%. While both contrasts taken together provided a significant contrast among the groups ($p < .06$), the second variate by itself did not ($p < .29$). Hence, only the first is considered further.

The magnitudes of the standardized coefficients of the first canonical variate indicated that the variables that contributed most to the differences, in order from most to least, are DIFB4 and DIFB1. The other variables contributed much less to the contrast. The signs of the coefficients for the

Table 46

Discriminant Analysis: 03-02 Difference Power Test Variables

First Canonical Variate - 3 Groups - 73.7% of Canonical Variation

	Discriminant Function Coefficients for Variables							Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-.5722	-.1878	.1825	-1.6386	-.0830	-.2682	-.1517	-1.04	-.19	1.27
Standardized	-.6340	-.2348	.2338	-.9274	-.1073	-.3063	-.1091	--	--	--

Second Canonical Variate - 3 Groups - 26.3% of Canonical Variation

Raw	.0712	.4462	-.7345	-.1069	-.3229	-.3472	-.0145	-1.03	.23	-.75
Standardized	.0800	.5578	-.9407	-.0605	-.4175	-.3966	-.0105	--	--	--

Group I vs. AS Canonical Variate

Raw	-.5595	-.1327	.0926	-1.6395	-.1213	-.3081	-.1523	-1.15	--	1.17
Standardized	-.6286	-.1658	.1186	-.9279	-.1569	-.3519	-.1096	--	--	--

Group AS vs. SA Canonical Variate

Raw	-.5142	-.4052	.5620	-1.2987	.1118	-.0282	-.1176	--	-.28	1.48
Standardized	-.5776	-.5065	.7198	-.7351	.1445	-.0323	-.0846	--	--	--

Table 46 (Continued)

Group AS vs. SA Canonical Variate

	Discriminant Function Coefficients for Variables							Function Values for Groups		
	DIFB1	DIFB2	DIFB3	DIFB4	DIFB5	DIFB6	DIFB7	AS	SA	I
Raw	-.2581	.2672	-.5100	-.9975	-.3148	-.4377	-.0962	-1.44	.08	--
Standardized	-.2900	.3340	-.6532	-.5646	-.4070	-.4999	-.0692	--	--	--

Bartlett's Chi Square Test of Significance

Variate(s)	χ^2	df	p
First and Second-3 Groups	23.23	14	.06*
Second-3 Groups	7.37	6	.29
I vs. AS	15.16	7	.04*
I vs. SA	11.28	7	.13#
AS vs. SA	8.55	7	.29

* Significant
# Approaching significance

variables DIFB4 and DIFB1 were both negative indicating but one factor in the discrimination among the groups. The values of the function for the groups AS, SA, and I, -1.04, -.19, and 1.27, respectively, indicating that the contrast is basically between group AS and I, though group SA is more like group AS than group I.

Overall groups univariate analysis: 03 - 02 difference power test variables. The significance of the differences among the three groups for each of the 03 - 02 variables DIFB1 through DIFB7 was tested by analysis of variance; the statistics for the univariate F tests are reported in Table 45. The differences were significant for the variables DIFB3 ($p < .06$) and DIFB4 ($p < .01$), and approached significance for the variable DIFB6 ($p < .13$).

Pairs of groups analysis: 03 - 02 difference power test variables. The statistics for the tests of the differences by pairs of groups are reported in Table 47. The multivariate differences between groups I and AS were significant ($p < .04$), and approached significance for the groups I and SA ($p < .14$). The multivariate differences between groups AS and SA were not significant. Univariate differences between groups I and AS were significant for the variables DIFB4 ($p < .003$) and DIFB6 ($p < .05$), and approached significance for the variable DIFB1 ($p < .11$). Univariate differences between groups I and SA were significant for the variable DIFB3 ($p < .04$), and approached significance for the variable DIFB4 ($p < .11$). While the multivariate differences between groups AS and SA were not significant, the univariate differences were significant for the variables DIFB3 ($p < .05$) and DIFB4 ($p < .07$).

Table 47

MANOVAs and ANOVAs for 03-02 Difference Power Test Variable by Pairs of Groups

I vs. AS

Variables	Hypothesis Mean Squares	F	p	df
DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6, DIFB7	(multivariate)	2.95	.04*	7/15 N/D
DIFB1	3.65	2.89	.11#	1/21 hyp/error
DIFB2	2.19	1.40	.25	1/21
DIFB3	.00	.00	n.s.	1/21
DIFB4	3.77	11.78	.003*	1/21
DIFB5	1.46	.87	.37	1/21
DIFB6	5.72	4.38	.05*	1/21
DIFB7	.10	.19	n.s.	1/21

I vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, D FB6 DIFB7	(multivariate)	1.94	.14#	7/15 N/D
DIFB1	2.77	2.19	.16	1/21 hyp/error
DIFB2	2.10	1.34	.26	1/21
DIFB3	8.23	5.01	.04*	1/21
DIFB4	.95	2.96	.11#	1/21
DIFB5	.02	.01	n.s.	1/21
DIFB6	.63	.48	n.s.	1/21
DIFB7	.10	.19	n.s.	1/21

Table 47 (Continued)

AS vs. SA

DIFB1, DIFB2, DIFB3, DIFB4, DIFB5, DIFB6 DIFB7	(multivariate)	1.35	.30	7/15 N/D
DIFB1	.13	.10	n.s.	1/21 hyp/error
DIFB2	.02	.01	n.s.	1/21
DIFB3	7.79	4.75	.05*	1/21
DIFB4	1.13	3.53	.07#	1/21
DIFB5	1.21	.72	n.s.	1/21
DIFB6	2.89	2.22	.16	1/21
DIFB7	.00	.00	n.s.	1/21

* Significant
\# Approaching significance

Summary of 03 - 02 Difference Power Test Variable Comparisons

The multivariate differences among the three groups for the 03 - 02 difference power test variables were found to be significant. The differences were found to be primarily between groups I and AS, though group SA was most like group AS. The variables that account for the differences were DIFB4, DIFB3, DIFB1, and DIFB6. These differences are fully accounted for by the changes in the group differences for these variables from observation 02 to 03.

Summary of the Data

The preceding sections report the data from each of the several observations along with the tests of the hypotheses to which the

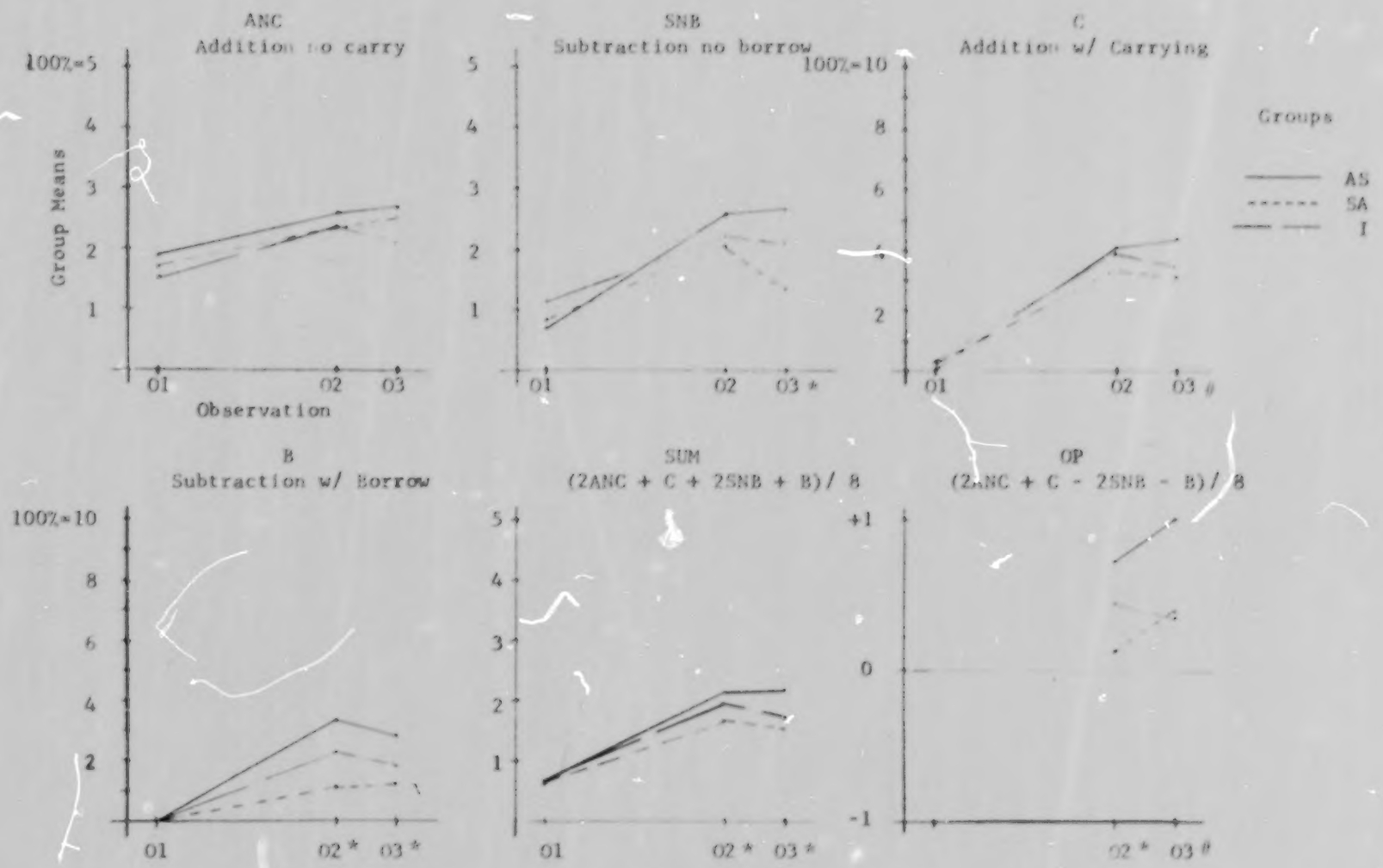
the data relate. Though detailed summaries of each set of data are found in the closing paragraphs of each of the preceding sections, the purpose of this section to provide a total picture of the findings of the various sections. To this end a summary of the multivariate and univariate tests of group differences for the data of observations 02, 03, and 03 - 02 is presented in Table 48; and since most of the differences found involve the basic achievement variables, Figures 9 and 10 provide graphical displays of the changes in performance for these variables from 01 to 02 and then from 02 to 03. The summaries follow the order of the preceding sections.

The observation 01. Baseline data of initial performance for the achievement variables was provided by observation 01. While mean scores at 01 for the variables B and C were essentially zero, the means for ANC and SNB were about 1.5 out of a possible 5.0. As expected, there were no overall group differences for any of the variables of observation 01.

The periodic observation data. The data from the 12 periodic observations were consolidated into data for four time periods to give profiles of group performance during instruction (see Figure 7). These data were related to the hypothesis that there were no group differences in the rate of learning (hypothesis H1). This hypothesis was broken down into several sub-hypotheses that were tested by the group means for the variables of the periodic observations; several of these were rejected. A summary of these tests is found in Table 13.

TABLE 48
p Values of F Tests Involving O2 and O3 Variables

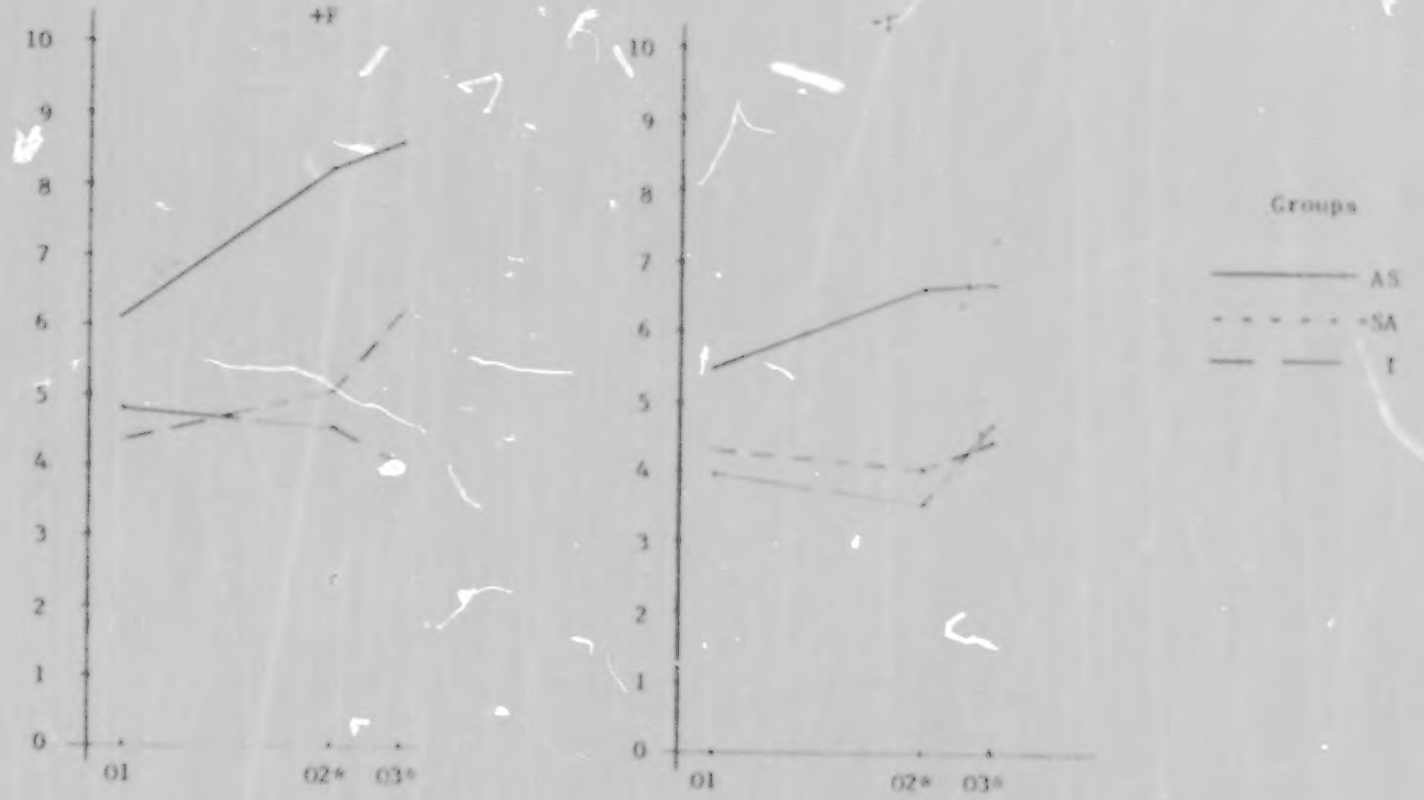
Variables	Source of Scores and Contrasts											
	O2				O3				O3 - O2			
	Gps	I,AS	I,SA	AS,SA	Gps	I,AS	I,SA	AS,SA	Gps	I,AS	I,SA	AS,SA
Six Achievement Variables	.03	.02	---	.01	.002	.003	.04	.02	.22	---	.08	.33
+F	.001	.001	---	.001	.001	.001	.01	.01	---	---	.12	---
-F	.001	.001	---	.01	.08	.08	---	.04	---	---	---	---
A:C	---	---	---	---	---	---	---	---	---	---	---	---
C	---	---	---	---	---	---	---	.15	---	---	---	---
SNB	---	---	---	---	.02	---	.07	.01	.13	---	.13	.06
B	.02	---	.09	.01	---	---	---	---	---	---	---	---
SUM	---	---	---	.08	.07	.11	---	.03	---	---	---	---
OP	.08	---	---	.03	---	---	.15	---	---	---	---	---
RGP	---	---	---	---	---	---	---	---	---	---	---	.13
INT	---	---	---	---	.01	---	.01	.02	.14	---	.08	.12
SUMB1 through SUMB7	---	---	---	---	---	---	---	---	---	---	.13	---
SUMB1	---	---	---	---	---	---	---	---	---	---	---	---
SUMB2	---	---	---	---	.15	---	.05	---	.09	---	.04	---
SUMB3	---	---	---	---	---	---	---	---	---	---	---	.07
SUMB4	---	---	---	---	---	---	---	---	---	---	.09	---
SUMB5	---	---	---	---	---	---	---	---	---	---	.07	---
SUMB6	.15	---	.06	---	---	---	---	---	---	---	---	---
SUMB7	---	---	---	---	---	---	---	---	---	---	.13	---
DIFB1 through DIFB7	.11	.06	---	.05	.10	.06	.13	---	.06	.04	.14	---
DIFB1	---	.10	---	---	.08	---	.04	.09	---	.11	---	---
DIFB2	.02	.01	---	.02	---	---	---	---	---	---	---	---
DIFB3	.01	.14	.07	.01	---	---	---	---	.06	---	---	---
DIFB4	.05	.02	---	.11	.09	.03	---	---	.01	.01	.11	.07
DIFB5	---	---	---	---	---	---	---	---	---	---	---	---
DIFB6	---	---	---	.13	---	---	---	---	.13	.05	---	---
DIFB7	---	---	---	---	---	---	---	---	---	---	---	---



* - Indicates that significant differences exist between groups.
\# - Indicates that marginally significant differences exist between groups

Graphs of Group Means for Some Achievement Variables

FIGURE 9



+F is the difficult addition facts measured under timed conditions at 01, 02, and 03.
-F is the difficult subtraction facts measured under timed conditions at 01, 02, and 03.

Graphs of Adjusted Group Means for +F and -F

FIGURE 10

Differential effects of the sequences on the rate of learning are most evident with respect to the variable B, borrowing. The mean of group SA was .24 after a three-week instructional period while group AS reached a level of .71 during a period of slightly less than three weeks. The rate of learning of group I was more gradual for variable B, but there is an interesting jump in performance from period T2 to T3 that is mentioned again later.

In general, however, the sequences facilitated learning in ways that reflected opportunity to learn. When the children of group AS were studying the addition algorithm they learned to carry, and when they studied the subtraction algorithm, they learned to borrow. Likewise, the means of group I for the four basic variables ANC, C, SNB, and B reflected instruction by increasing with each succeeding time period. The means of group SA, however, while also reflecting the instructional treatment, were in general less than those of either group AS or I.

In addition to the group differences just noted, the data for the four time periods exhibited two other patterns. First, the jump in the performance of group I mentioned earlier was not restricted to variable B but was a part of a general increase in the means of group I for all of the variables. This time period corresponds with the introduction of compact notation algorithms into the learning activities of group I. The performance of group SA for the variable B was zero prior to the introduction of compact notation into their

learning activities while the performance of group AS for the variable B increased sharply at P10 which was just prior to the introduction of the compact notation algorithm (see Tables 4 and 10).

The second pattern evident in the profile data is that with a single exception the means of each group for the basic variables were ordered from greatest to least as ANC, SNB, C, and B. The exception was for group SA during the second time period where the means for C and B were reversed, however, both were nearly zero.

The posttest data: observation 02. The observation 02 measured three sets of variables: the achievement variables, the total power test variables, and the difference power test variables. These data related to the hypotheses H2, H3, and H4, respectively. While hypothesis H2.02 was rejected, and some evidence was found of group differences for the difference power test variables, no evidence was found to substantiate group differences for the total power test variables.

The means for the achievement variables of observation 02 revealed the following patterns of group differences. The means of group AS exceeded those of the other groups for all six of the directly measured achievement variables. With exceptions involving the variables +F and -F, the means of group SA were the least for these six variables. The discriminant analysis of the group means at 02 indicated that the multivariate differences involved a contrast of group AS with the both groups SA and I; groups SA and I could not be distinguished

from each other in the multivariate space (see Table 20). The multivariate and univariate tests of differences between pairs of groups further indicated that group AS was significantly different from both groups SA and I; these two groups differed significantly only for the variable B.

The retention test data: observation 03. The observation 03 was simply a repeated administration of the tests of observation 02, and hence, also relates to the hypotheses H2, H3, and H4. Hypotheses H2.03 and H4.03 were rejected by the multivariate tests of overall group differences for the total power test variables.

The means of the groups for the 03 achievement variables revealed patterns of differences very similar to those found at 02, except that at 03 the differences were found to involve all three groups. The means of group AS for the six directly measured achievement variables again exceeded those of the other two groups. With exceptions involving the variables +F and ANC group SA again had the least means for these six variables. The discriminant analysis of the group means yielded two significant contrasts among the group means; the first that accounted for about 72% of the canonical variation, was seen to be a contrast of group AS with both groups I and SA, while the second was a contrast of groups AS and I with group SA. The multivariate and univariate tests of differences between pairs of groups further indicated that all three groups differed significantly from one another; the order of the means

from greatest to least was usually AS first, then I and finally group SA. Exceptions to this pattern were found for the variables ANC and +F where the means of group SA exceeded those of group I.

The group differences for the difference power test variables of 03 indicated that the means of group I may have been more uniform with respect to addition and subtraction than those of group AS, though the meaning of the differences was not clear.

The rate of forgetting: observation 03 - 02. The data for observation 03 - 02 were constructed from the data of 02 and 03, and the tests provided by this data therefore are not independent of the tests already summarized. They do provide direct tests of group differences in the rate of forgetting. The data from 03 - 02 can also be related to hypotheses H2, H3, and H4. Only the hypothesis H4.03 - 02 was rejected by the multivariate tests of overall group differences. Further tests of group differences indicated that the group differences for the changes in the difference power test variables were fully accounted for by the data from observations 02 and 03. Though hypotheses H2.03 - 02 and H3.03 - 02 were not rejected, there was some evidence of differences in the rates of forgetting between groups I and SA for some of the achievement and power test variables.

Chapter IX

CONCLUSIONS

The Findings of the Study

Recall that the purpose of this study was to investigate the differential effects on achievement of the three instructional sequences AS, I, and SA. Two questions whose answers accomplish this purpose are stated on pages 43, 44 in Chapter V. The data collected to answer these questions are reported in Chapter VIII. The answers provided by these data, and therefore the conclusions of this study are presented in this chapter.

Findings Related to Achievement Monitoring Data

The data from the twelve observations of group performance made during the course of instruction were consolidated into data for four time periods; these data provided the basis for some statistical tests of group differences. Since these measures were not direct measures of learning rate, no direct conclusions about the rate of learning can be made. However, some implications about learning rate based upon these data are considered later in this chapter. The tests of group differences for each of the time periods together with the general form of the achievement monitoring curves warrant the following conclusions.

1. The learning profiles reflected opportunity to learn in a direct way, but the instructional sequence SA produced generally lower performance than that of the other two groups. Furthermore, the tendency of all groups was to learn addition first no matter what instructional sequence was being used.

Findings Related to Observations 02 and 03

The effects of the sequences on the amount of learning and the retention of learning were measured by achievement and power test variables at observations 02 and 03. The analysis of these (see Chapter VIII) leads to the following conclusions.

2. The amount of learning and retention measured by the achievement variables was significantly greater for the sequence AS than for either of the other two sequences.

3. Excluding performance for the addition facts and addition that requires no regrouping, the overall performance of group I for the achievement variables was significantly greater than that of group SA at the end of the retention period. Non-significant differences of smaller magnitudes also favored group I over group SA immediately following instruction.

The significant differences among the group means for the power test variables involved only the difference measures for similar

aspects of the addition and subtraction algorithms. While these differences are generally difficult to interpret, the following conclusion seems indicated.

4. Following instruction the power test performance of group AS for subtraction was more similar to their performance for addition than was true of the other two groups. At the end of the retention period, however, the power test performance of group I was the most uniform.

Observation 03 provided the basic data for comparisons of retained learning. However, mean differences from observation 02 to observation 03 provided direct measures of the rate of forgetting. Significant overall differences among the groups in their rate of forgetting were found only for the difference power test variables DIFB1-DIFB7. These differences do not follow any clear pattern except to account for adjustments in mean scores for these variables from 02 to 03. The conclusion related to 03 - 02 data, therefore, is as follows.

5. Differences in the rate of forgetting associated with the three sequences are found only for the power test measures of operation differences and are not regarded as important.

The facts performance question of this study was, "Do measures of proficiency with the addition and subtraction facts show significant changes due to the incidental drill inherent in learning the

usual algorithms?" The data that bear on this question are best summarized by the graphs of Figure 10 (see Chapter VIII, page From observation 01 to observation 02 the means of group AS for the variables +F and -F increased, while the means of group I decreased and the means of group SA increased for +F and decreased for -F. Curiously slight increases were recorded for these variables from 02 to 03 with the single exception of the mean of group I for addition facts. However, as is apparent from Figure 10 and the distributions of scores for +F and -F, the increases and decreases were quite small in comparison to the total range of the scores. Thus, one conclusion is as follows.

6. Instruction with the addition and subtraction algorithms did not result in increases in performance on addition and subtraction number facts.

Some implications of the preceding six conclusions are considered following a discussion of the limitations that apply to them.

The conceptualized study was to include measures of IQ by treatment interaction and was to attempt to validate the behavioral hierarchy from which the sequences of behaviors were selected. However, the number of subjects that were available was too small to allow for a meaningful test of these interactions, and the difficulty of relating the paper and pencil test items to the stated behaviors made it seem wise to defer the attempts to validate the behavioral hierarchy to a later time.

Limitations

Any study is subject to factors that limit the findings and implications of that study. Many factors of this type have been identified by Campbell and Stanley (1963). Factors of the present study that may limit the findings or implications are discussed in this section in terms of the quality of the instructional materials, the tests used, the analysis, the confounding factors, and the generalizability of the findings.

The Quality of the Instructional Materials

The instructional materials were to accommodate three sequences of behavioral objectives. Though the sequences were chosen in ways consistent with the postulated dependencies of the given hierarchy, the hierarchy itself was not empirically validated. Detailed methods for validating a hierarchy have been proposed, but these require valid measures of each behavior of the hierarchy (see Eisenberg and Walbesser, 1971). And since the variables measured in this study cannot be properly called valid measures of the behaviors of the hierarchy used, data from this study were not used to test the validity of the hierarchy. Heimer (1969) has pointed out, it is not logically sound to test hypotheses about learning hierarchies and maintain the option of deciding whether the hierarchy was valid or not. Therefore, it is necessary to identify those aspects of the hierarchy that are assumed to be valid.

It has been assumed that the specification of the behaviors is sufficient and that the tasks A and S, as have been defined in Chapter V, are independent. However, if certain necessary behaviors were not included in the hierarchy, and hence not taught in the three treatments, the expected results would be generally low performance levels for the terminal objectives of the hierarchy. And, if performance were sufficiently low, real group differences may not be detected by the post instruction tests of performance. If the two tasks A and S are not independent, the findings of this study cannot be generalized to a hierarchy of different behaviors with tasks that are independent. Though these factors must be considered, no logical reasons are seen to deny the sufficiency of the hierarchy or the independence of the tasks A and S.

The Ss, however, did not interact with the sequences directly, but with activities that were designed to embody the sequences. The findings of the study, therefore, are also limited by the adequacy and integrity of the activities. The essential considerations are that the instructional activities do reflect the sequences specified, and that children do learn from the activities.

The activities were developed to fit the requirements of the three sequences described in Chapter VI. Many of the activities that were used in the present study, had been used earlier by children of the appropriate ages and experience levels, and seemed to

to be pedagogically sound. These factors, together with frequent observations of the classes and even more frequent consultations with the teachers, support the claim for the integrity of the treatments and the expectation of adequate achievement levels. The effects of refining the hierarchy or modifying the activities cannot be predicted, however, the quality of the materials seems adequate to provide a fair comparison of the effects of using the sequences AS, SA and I.

Limitations of the Tests

In addition to limitations imposed by the materials used, there are limitations of instrumentation imposed by the tests used. These involve the dependent variables that were chosen, and the items that were used to measure performance for those variables.

It is always desirable to obtain a more complete description of outcomes; variables of interest cannot always be anticipated or measured in any efficient manner. In particular, the following variables were not measured: power test measures of performance with the addition and subtraction facts before and after instruction, measures of transfer, and observations of the methods the children used to produce their responses. While performance two weeks after instruction was examined, it would be interesting to know the retention levels of behaviors over longer periods of time. In short, different

or additional selections of variables might yield a modified picture of the effects of the three sequences.

Test validity and reliability can also be limiting factors in a study. However, the test items used can be characterized as random selections from finite populations that are defined by the variables; the tests therefore have content validity. Though different selections from the item populations might produce different results, this does not seem likely. Furthermore, since the Hoyt reliabilities for the variables (reported in Appendix C) range from .70 to .92, no limitations to the findings are attributable to either test validity or reliability.

The above limitations relating to choice of variables, and the items used to measure performance are cited as possible sources of limitations in the tests. While one should be aware of these factors, the judgment is made that the quality and quantity of evidence obtained are adequate for the conclusions of the preceding section.

Limitations of the Analysis

The findings of an experimental study are limited not only by the integrity of the treatments and the adequacy of the criterion measures, but also by the appropriateness of the analysis of the data. The difficulty in the present study concerns the unit of analysis.

The basic unit of analysis was taken as the IQ block. There were 18 IQ blocks with not more than three subjects in each block; from these blocks the subjects were randomly assigned to the three treatments. The analysis model based upon these randomized blocks requires that cells be treated independently. However, the cells were not treated independently. There were only three treatment groups; furthermore, the members of a group were encouraged to interact with one another in the presence of their teacher. However, if the unit of analysis is taken to be the entire treatment group, no statistical tests are possible.

The use of IQ blocks as the unit of investigation may be regarded as a compromise between using individuals, where there are many more possible dependencies, and using groups where no tests are possible. The effect of this compromise can be analyzed in terms of the power of the statistical tests employed. That is, the use of the IQ blocks requires 18 degrees of freedom as opposed to the single degree of freedom that would be required in an analysis of covariance using IQ as a covariate. Thus, the use of IQ blocks greatly reduces the power of the tests and is an attempt to be conservative. The precise effect of violating this assumption is not known.

Confounding Factors

The role of the teacher. An important factor in any instructional treatment is the role of the teacher; as noted earlier (see Chapter VII),

the teachers are in reality part of the treatments. This fact contributes to the lack of independence between treatment blocks discussed earlier and also raises the alternate hypothesis that differences result from teacher effects with little regard for the treatments.

Reasons for rejecting the hypothesis of differential teacher effects are as follows:

1. The teachers were randomly assigned and reassigned to groups.
2. The learning activities specified what the children and the teacher were to do.
3. The role of the teacher was that of an administrator of a given set of activities.
4. The teachers were experienced with children of these ages and materials of the type used.
5. The experimenter was present in the school for 16 of the 26 days of the instructional period, visiting classrooms and consulting with teachers on a regular basis.

There was no apparent reason why any teacher should have been superior to any other. They were all cooperative and careful to follow both the descriptions and intent of the activities.

The above discussion suggests that it is not possible to predict how the outcomes of this investigation would have been affected by

teacher differences. There is strong evidence, however, in the periodic observation data (see Chapter VIII) that the sequences affected learning in some entirely predictable ways. Thus, while it cannot be demonstrated that no differential effects were present in the results of this study, the hypothesis that all effects are due to teacher differences is clearly untenable.

Limitations of History. Any study conducted in an operating elementary school encounters events for which no adequate control exists. Events that could affect the outcomes of this study have been identified and discussed in the last section of Chapter VII. While it is not clear if these events are threats to the internal or the external validity of the study, they should be noted as possible limitations to the findings. Without repeating the discussion of Chapter VII, these events consist of an outbreak of mumps and measles that affected attendance during the first three weeks, and the absence of the teacher assigned to group SA on the second, fourth, fifth and sixth instructional day because of illness.

While the absences of children can be regarded as random across groups and so of no real importance to group comparisons at O2 and O3, the effects of teacher absence are more problematical. The absences of their teacher could normally be expected to affect the achievement of group SA adversely. The following factors, however, are cited to support the contention that these absences had no real effect on the outcomes of the experiment:

1. The teacher aide, who had also worked extensively with the children and with instructional materials of the kind used in the experiment, assumed a major role during the absence of the teacher.
2. The other teachers, the teacher aide, and the experimenter met with the two substitute teachers to discuss the work to be done.
3. Since the absences occurred during the first six days of instruction, the bulk of instruction, 20 days out of 26, followed the time of the last absence. Furthermore, during the first six days all of the children were adjusting to altered classoom and teacher arrangements due to the random assignment of subjects to treatments.

Because of the reasons listed above, it is assumed that these teacher absences had no effect on the outcomes observed at 02 and 03. However, if an effect does exist, the result of it should be to decrease the performance of group SA particularly with respect to subtraction.

Limitations to Generalizability

The factors that have been discussed relate primarily to the findings of the study. The following paragraphs present factors that limit the application or generalizability of the findings. The factors considered here are the type of instructional materials employed and the particular population selected.

The Type of Materials Used

A general description of the type of instructional activities employed in this study is given in Chapter V and will not be repeated; however, it should be noted that the results were obtained using a particular kind of instructional approach. The findings may not generalize to situations where the instructional methods or materials are basically different from those employed in this study.

The Population

While a complete description of the subjects is presented in Chapter VII, the summary given here identifies attributes that could affect the generalizability of the findings. The subjects were lower grade elementary school children who were able to respond correctly to addition and subtraction facts, and who were unable to use an addition or subtraction algorithm for two-digit whole numbers beyond simply counting. They were also familiar with the type of instructional materials used. Additionally, however, the subjects were also from an inner city neighborhood school. This neighborhood had been the focus of a community redevelopment project and consequently was not experiencing the very high turnover rates of school population and staff that is frequently associated with inner city schools. The morale of staff and students seemed to be very good, and the class size for mathematics in the lower grades at the time of the study was about 20 students. Though these factors are cited, the judgment is made that the findings do apply to populations that do not have all of

the above characteristics. The characteristics of primary importance are seen to be those of age and entering behaviors.

Discussion of Results and Implications for Sequencing

If the conclusions are valid then sequence I was not as productive as sequence AS and was more productive than sequence SA. It should be kept in mind, however, that the differences were only apparent for the achievement measures. While the power test measures of performance gave some indications of group differences, these differences were small and the meaning was not clear.

The results of comparing sequences I and AS is entirely consistent with the results of an earlier investigation (Wiles, Romberg, and Moser, 1972) that used the same sequences but different activities (see Chapter III). Furthermore, the subjects of the earlier investigation were similar to those of the present investigation in terms of entering behaviors and age, but they were quite different in terms of socioeconomic status. Hence the finding of greater performance for sequence AS has some robustness. However, since the sequence SA, with few exceptions, produced lower performance than that of sequence I, a simple comparison of non-integrated versus integrated sequencing will not account for the results of this study.

In Chapter III it was noted that the concept learning evidence indicates integrated sequencing strategies will facilitate the learning of coordinate concepts, but that the integrated sequencing strategy

had not proved to be advantageous for algorithmic learning. However, as was pointed out in Chapter I, the usual addition and subtraction algorithms are not simple concept identification or concept usage tasks, but tasks that involve chains of behaviors that presuppose the acquisition and use of many concepts.

The results of the present investigation together with the research related to concept learning, indicate that the complexity of the tasks must be considered along with the structural similarities of the tasks. The following paragraphs attempt to account for the findings of this study in terms of the relative difficulty of the tasks A and S, of facilitative transfer, and of the amount of focused practice provided by the three sequences. Implications for the sequencing of coordinate tasks based upon this account are then stated.

First, it should be noted that the mean scores at observation 02 and 03 along with the profile data indicate that the addition algorithm is more easily acquired than is the subtraction algorithm. In particular, the group means for the variable OP, which compare addition and subtraction achievement, were never negative. Though most of the means of group AS for the power test variables DIFB1 through DIFB7 were negative at observation 02, the absolute values were small; furthermore, the means for these variables were positive at 03. Thus, the relative difficulty of addition and subtraction algorithms, as determined by this study, is consistent with the findings of Wiles, Romberg and Moser (1972), and of Romberg and Planert (1970).

Secondly, the profile data provide evidence that the acquisition of task A behaviors facilitated the acquisition of task S behaviors. The important difference between the groups AS and SA with respect to subtraction was that the performance of group AS for the variable C, carrying, was at a high level before they began to study subtraction, while the mean of group SA for the variable C was practically zero. The profile data show that the mean of group AS for the variable B, borrowing, increased much more sharply during instruction for subtraction and reached higher levels than the mean of group SA. Similarly, the mean of group I for the variable B increased sharply after performance for the variable C had reached a high plateau. While one might expect a similar facilitative effect from task S to task A, the data provide no evidence that this occurred. The only opportunities to have observed such evidence was in the data for groups SA and I. The children of group I acquired carrying first and so disallowed any test, and the performance of group SA was not at a high enough level to provide a valid test of transfer from task S to task A.

Two factors now seem to be evident. The first is that acquiring the less difficult behaviors associated with task A facilitated the acquisition of behaviors associated with task S; the second is that before one of the tasks, in this case A, is acquired, any transfer effects are likely to be small. Stated simply, structural similarities between complex coordinate tasks may promote transfer from

one to another, provided that the first is acquired and not just taught. An optimal sequencing strategy then that relies upon transfer from one task to the next requires a sequence that provides for early success in acquiring some one of the tasks.

Sequencing for early success has another advantage apart from providing the possibility of transfer to another task. Simply, it is that a task cannot be practiced before some degree of mastery in doing the task is achieved. This is illustrated by the data of this study in the following way.

The performance of group AS exceeded that of groups SA and I for all of the achievement variables at observations 02 and 03. Included among the achievement variables are the variables relating to facts performance, +F and -F. While the overall increases in facts performance from 01 to 02 as measured by the variables +F and -F were trivial, there were significant differences among the group means at 02 ($p < .03$) and 03 ($p < .002$) for these variables; the means of group AS accounted for two of the increases in performance for these variables and group SA the third. The increase of group SA was for the variable +F, and the only real successes of group SA were with respect to addition. Now since the only practice with the facts was provided in the context of solving two-digit problems, it is evident that the amount of practice was dependent upon the ability to complete the task. Given success, the isolating of the tasks A and S for groups AS and SA served to concentrate the practice and so

increase the performance of groups AS and SA for the variables +F and -F even further. What appears to be true, therefore, is that sequence AS provided success in acquiring the tasks, the prerequisite for practice, along with concentrated practice with the tasks. The resulting means were significantly greater than those of the other two groups. Sequences SA and I on the other hand did not provide the early successes of group AS, and sequence I never provided focused practice.

What has just been outlined is not only an explanation for the increases of group AS for the variables +F and -F, but also a further explanation of the higher performance of group AS for the achievement variables in general. It is simply that the earlier success and concentrated practice provided by sequence AS resulted in higher performance for the achievement variables.

Now while the two factors of success and focused practice can account for observed differences during the course of learning and the order of the group means for the achievement variables following learning these factors do not explain decreases in performance from 01 to 02 for the variables +F and -F by group I and for -F by group SA. However, information provided by the children's test sheets does provide an explanation for this.

The following examples are copies from the children's tests.

$$\begin{array}{r} {}^{1}6 \\ +\ 8 \\ \hline 14 \end{array} \quad ; \quad \begin{array}{r} {}^{0}1 \\ \times 1 \\ \hline 01 \end{array} \quad , \quad 15 - 9 = 6 \quad , \quad \begin{array}{r} {}^{1}6 \\ +\ 4 \\ \hline 10 \end{array} \quad ; \quad \begin{array}{r} 15 \\ -\ 6 \\ \hline 19 \end{array}$$

These examples are merely illustrative; many more such items could have been chosen. Table 49 indicates the proportion of children making responses that showed the influence of correct or incorrect algorithmic behavior. This involves nearly one-half of the children and the proportion is probably low since a child was not counted unless it was clear that he was processing facts as in the two-digit algorithm; simple reversals such as $\begin{array}{r} 15 \\ -\ 9 \\ \hline 14 \end{array}$ or mixed operations such as $9 - 4 = 13$ were not counted at 02 or 03.

Table 49

Proportion of Facts Tests Showing Algorithmic Interference

Group	Observation 01	02	03
AS	2/15	7/13	5/13
SA	0	8/16	8/16
I	0	6/15	5/14
TOTAL	2/49	21/44	18/43

It is apparent from the tabulation that this behavior was generally

present in all groups and at both observations 02 and 03. That this behavior was induced by the treatments is demonstrated by the fact that the only two examples on papers from observation 01 that can be related to such algorithmic interference were $-\frac{15}{11}\frac{6}{}$ and $-\frac{10}{10}\frac{5}{}$, neither one of which would have been considered adequate evidence of algorithmic interference at 02 or 03.

The reason for decreases in facts performance from 01 to 02 now seems to be clear. Children in all of the groups experienced "algorithmic interference" that undoubtedly decreased their scores for +F and -F. The interference was enough to counterbalance any increases in performance due to incidental drill except in the cases of the more focused drill provided by sequences AS and SA.

Summary

The observed results of using the three sequences for the tasks A and S can now be accounted for as follows. The sequence AS provided almost immediate success for addition and hence focused practice. Learning the addition algorithm had a facilitative effect on learning the subtraction algorithm and so sequence AS provided focused practice for subtraction as well. Sequence SA did provide success with task A, however, and along with this success came focused practice for the addition behaviors. Sequence I provided success with addition later than did sequence AS and success with subtraction very little before sequence AS. Though there was some evidence of transfer from task A to task S in the performance of group I, the only possibility of more

practice than provided by sequence AS was in subtraction, but this was never focused practice. It may be concluded that the greater opportunity for transfer and practice provided by the earlier success and concentrated practice of sequence AS raised the performance of this group significantly above that of the other two groups.

The effects of instruction for the algorithms on facts performance includes evidence of small increases when the algorithms are learned as the practice is focused; but there is stronger evidence of algorithmic interference in the measures of facts performance. Since the increases where they occurred were quite small, and the performance levels actually decreased for half of the cases, a reasonable conclusion is that facts performance was hindered as much as it was helped by the instruction for the algorithms. The incidental drill hypothesis has very little support in the findings of this study.

Implications for sequencing structurally similar coordinate tasks of complexity similar to that of tasks A and S can now be stated.

The tasks should be sequenced so as to provide the earliest possible success, that is, the easier task should be presented first. This does not necessarily mean that the other task should be ignored, as in the present study, while the first task is being learned. It does mean that the primary focus of instruction should be on the easier task first. This implication identifies the relative difficulty

of the tasks as another dimension of a behavioral hierarchy for learning that needs to be considered. Resnick (1970) in a discussion of issues relating to learning hierarchies advocates this kind of analysis for sequencing as follows.

> Often, two superficially similar tasks differ with respect to their demands on some basic function such as memory or perceptual organization. These differences between tasks, together with information concerning the shared components which would facilitate transfer, provide the bases for ordering tasks according to complexity and thus for predicting optimal instructional sequences (p. 6).

Resnick continues by saying, "It should be noted that this method of analysis into components is limited to behaviors that can be characterized as 'chains' or 'procedures' (p. 6)." This discussion is in line with the implications of the study.

Since care must be taken to avoid over-generalizing these implications, the following two limitations are noted.

1. If the tasks are easily learned by the population for which the instruction is intended, an integrated approach may work very well, as other research has shown (e.g., Short and Haughey, 1967). An additional example comes from the present study. In this study it is quite apparent from the periodic observation data that addition and subtraction not requiring regrouping are readily acquired with no specific instruction; the more complex regrouping behaviors caused the difficulty.

2. There may be reasons for choosing a sequence that is less efficient in terms of immediate acquisition of the behaviors. A

whole range of philosophical, psychological and pedagogical considerations must be made when an optimal instructional sequence is considered in broader terms than just efficient acquisition of the behaviors.

Implications for Research

The questions and issues raised by research may be as important as the results of research. The outcomes of this study bear on several issues that could lead to further investigations. These issues concern the identification of coordinate tasks, the specification of tasks and dependent variables, and the learning of algorithms.

The Identification of Coordinate Tasks

To what extent is the independence of tasks an empirical question and to what extent is it a logical question? This issue is related to hierarchy validation as outlined in the section on limitations presented earlier in this chapter. This issue arises in the following way.

For the present study, the coordinate nature of the tasks A and S was decided on a logical basis. Since there seemed to be no reasons to require that one task be prerequisite for the other, the tasks A and S were assumed to be coordinate. However, the obtained data are consistent with the hypothesis that the tasks are not coordinate.

In particular, if the behaviors of task A were prerequisite for those of task S, one would expect that the relative addition performance of each group would exceed their subtraction performance. This

was exactly what was observed at every observation; the variable OP was never negative. Furthermore, if task S were dependent upon task A one would expect the sequences SA and I to be less efficient than sequence AS. And, since sequence I provides opportunities to acquire the behaviors of both tasks throughout instruction, while sequence SA requires children to attempt to acquire the behaviors of task S before they have been given instruction for the "prerequisite" behaviors of task A, sequence I should produce better performance than sequence SA; this also was observed.

Though these data are consistent with the hypothesis that task S is dependent upon task A, they do not establish such a dependency; they can be accounted for in another way that has nothing to do with task dependency. Simply stated, the behaviors of task A are more easily acquired than are the behaviors of task S. This latter explanation furthermore is consistent with the dependencies of the learning hierarchy produced by a logical analysis of the addition and subtraction algorithms for whole numbers; the hierarchy does not postulate any dependencies between the tasks A and S.

Methods proposed for empirically validating logically derived learning hierarchies deal with hypothesized dependencies among the tasks rather than hypotheses of task independence (see Walbesser and Eisenberg, 1972). While an adequate treatment of how this has been done is not possible here, it can be noted that these procedures

usually involve some combination of logical and empirical methods. At any event the relationship between a logical and an empirical basis for testing the independence of tasks requires exploration.

The Specification of Tasks and Variables

There is a need for the careful specification of tasks and variables in research relating to the sequencing of the tasks for instruction. Of particular importance are the complexity of the task or tasks to be learned and the existence of relationships among the tasks that are to be learned. It would be useful to know if the tasks require single or multiple responses, if they were considered to be skills, concepts, procedural chains, processes, and so on. Where more than one task is to be learned, relationships among the tasks to be considered should include whether they are considered to be dependent, coordinate, or illustrative. Anything known about the relative difficulty of the tasks for the anticipated learners should also be specified.

In addition, a reasonably complete characterization of the tasks, the choice of dependent variables should include as many of the following as possible:

1. The rate and level of acquisition of each of the tasks to be learned during instruction.
2. Measures of the power, speed, and retention of the tasks after learning.
3. The effect of the instruction on the power and speed of performance for prerequisite tasks.

The Learning of Algorithms

This study also raises some questions about the learning of algorithms. The first of these concerns the use of expanded notation. In particular, what is the role of expanded notation algorithms in the learning of compact notation algorithms? This author suspects that if one is generally interested in learning algorithms, and not in just a few selected algorithms, expanded notation may play an important role as another example of how to do the same job, and thus might lead to the ability to construct algorithms.

Expanded notation may serve as a catalyst in enabling the learner to pass from a physical or pictorial representation to the symbolic representation. And as such the expanded notation forms may never be fixed in the learners memory or are soon forgotten in favor of more efficient ways of symbolizing the process. However, the hypothesized facilitative effect of expanded notation has not been demonstrated. Does expanded notation really facilitate or does it hinder the acquisition of compact notation numeration or computation? It may be argued that expanded notation deals with understanding. If this is so, it should be possible to specify some results of this understanding in terms of retention, validation, applications or transfer.

A second question related to algorithmic learning concerns the component skills or concepts that make up an algorithm. The existence

of many rather expensive drill and practice programs for all basic arithmetic skills, suggests that speed of response is thought to be very important. While it seems clear that some minimum response rate with the basic facts is necessary it seems reasonable to expect that there is a response rate beyond which greater speed has no effect. Most generally the question can be stated as, "How does speed in doing component tasks, or "sub-algorithms" affect the learning of new algorithms that incorporate them?"

A final question concerns the relative difficulty of learning addition and subtraction algorithms. The subtraction algorithm seems to possess some constructs that the addition algorithm does not. In particular, when one adds following the usual addition algorithm, the sum of a column is formed first and then the decision concerns what to do about the obtained sum; the simple rule, "Never write more than one digit in a column." handles the matter quite nicely. However, in the usual subtraction algorithm one first asks, "What will happen if I try to subtract in this column?" If the subtraction is not possible, a suitable subtrahend must be constructed. The processes involved in making the decision to regroup, and carrying out the regrouping, once it is decided to be necessary, seem to be different. Those processes involved with subtraction certainly require the reversability characteristic of the concrete operations stage of Piaget, and perhaps even elements of the formal operations

stage (see Flavel, 1963). While we do know it is quite possible to get children, even young children to "do an algorithm" it seems reasonable to ask just how developmental level relates to the elusive variable "understanding." The general question then is as follows: "What are the constructs of algorithms that affect their relative difficulty, and what are the relationships of these constructs to the developmental stage of the learner?'

Concluding Remarks

The research reported in this thesis was undertaken in response to questions about sequencing instruction to facilitate algorithmic learning. Of the three sequences employed, the sequence, addition then subtraction, was found to promote higher achievement on timed variables than the other two sequences, while the integrated sequence was somewhat more successful than the subtraction first sequence. These differences were accounted for in terms of the relative difficulty of the two algorithms, facilitative transfer, and focused practice. The limitations of the findings and implications involve the subjects used, the type of tasks to be learned, the kind of learning materials used, and the tests used.

The findings of this study suggest that in the absence of compelling reasons to do otherwise, structurally similar coordinate tasks of a complexity like that of the tasks of this study, should be sequenced consecutively with the easier coming first. The study raises several issues related to research dealing with sequencing variables

and algorithmic learning. These include the specification of tasks and variables, the role of expanded notation, and the effects of speed in performing prerequisite tasks on the tasks to be learned. The need for continued research into algorithmic learning is evident.

References

Baker, R. L. and Schutz, R. E. Instructional Product Development. New York: Van Nostrand, 1971.

Bartlett, M. S. Multivariate Analysis. Journal of the Royal Statistical Society (Suppl.), 1947, 9, 176-197.

Bock, R. D. Contributions of Multivariate Experimental Designs to Educational Research. In R. B. Cattell (Ed.). Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally, 1966.

Brownell, W. A. Learning as Reorganization. Duke University Research Studies in Education, No. 3, Durham, North Carolina: Duke University Press, 1939.

Brownell, W. A. and Moser, H. E. Meaningful Versus Mechanical Learning: A Study in Grade III Subtraction. Duke University Research Studies in Education No. 8, Durham, N. C.: Duke University Press, 1949.

Campbell, D. T. and Stanley, J. C. Experimental and Quasi-experimental Designs for Research on Teaching. In Gage, N. L. (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Cochran, W. G. Analysis of Covariance: Its Nature and Uses. Biometrics, 1957, 13 (3), 261-281.

Cronbach, L. J. and Snow, R. E. Final Report: Individual Differences in Learning Ability as a Function of Instructional Variables. Stanford: Stanford University, 1969.

Eisenberg, T. A. and Walbesser, H. H. Learning Hierarchies-Numerical Considerations. Journal for Research in Mathematics Education, 1971, 2(4), 244-256.

Ekman, L. G. A Comparison of the Effectiveness of Different Approaches on the Learning of Addition and Subtraction Algorithms in the Third Grade. (Doctoral dissertation, University of Minnesota) Dissertation Abstracts A, 1967, 27, 2275-2276.

Evans, J. L., Homme, L. E., and Glaser, R. The Ruleg System for the Construction of Programmed Learning Sequences. Journal of Educational Research, 1962, 55, 513-518.

Farris, D. C. A Study of Related Instructional Objectives Classified by Modes of Representation. Paper presented at the 49th annual meeting of the National Council of Teachers of Mathematics, Anaheim, April 1971.

Flavell, J. H. The Developmental Psycnology of Jean Piaget. Princeton: D. Van Nostrand, 1963.

Gagné, R. M. Learning Hierarchies. Educational Psychologist, 1968, VI(1).

Gray, W. L. The Effects of an Integrated Learning Sequence on the Acquisition and Retention of Mathematics and Science Behaviors in Grade Five. (Doctoral dissertation, University of Maryland) Ann Arbor, Michigan: University Micro-films, 1970 No. 70-22, 550.

Harvey, J. G., Green, M. I. and McLeod, D. B. The Task Analysis for Development Mathematical Processes, Arithmetic, Book 3: Writing Mathematical Sentences Using Larger Numbers and Arithmetic.

Book 4: The Integers. Working Paper No. 100. Madison: Wisconsin Research and Development Center, University of Wisconsin, 1972.

Harvey, J. G., Romberg, T. A., and Fletcher, H. J. Analysis of Mathematics Instruction: A Discussion and Interim Report. A paper presented at the First International Congress on Mathematical Education, Lyon, France, August 1969.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart and Winston, 1965.

Heimer, R. T. Conditions of Learning in Mathematics: Sequence Theory Development. Review of Educational Research, 1969, 39(4), 493-508.

Hickey, A. E. and Newton, J. M. The Logical Basis of Teaching: I. The Effect of Subconcept Sequence on Learning. Newburyport, Mass.: ENTELEK INC. 1964.

Jerman, M. Some Strategies for Solving Simple Multiplication Combinations. Journal for Research in Mathematics Education, 1970, 1(2), 95-128.

Kersch, Y. Engineering Instructional Sequences for the Mathematics Classroom. Research in Mathematics Education, Washington: National Council of Teachers of Mathematics, 1967.

King, B. W. On Scrambling Instructional Stimuli. Journal for Research in Mathematics Education, 1970, 1(4), 233-240.

Metropolitan Achievement Test, Primary II, Form F. Harcourt, Brace and Jovanovich, 1970. New York: (see Burroughs Mental Measurements Yearbook, V. 6, Entry 15.)

Murray, J. Studies in Arithemtic, Vol. I. Scottish Council for Research in Education, London: University of London Press, 1939.

Nelson, B. A. and Grayer, D. A. The Effects on Short- and Long-Term Retention of Discovery and Expository Methods of Presenting Selected Geometry Concepts: A Replication. Technical Report #208 of the Wisconsin Research and Development Center for Cognitive Learning. Madison: The University of Wisconsin, 1972.

Newton, J. M. and Hickey, A. E. Sequence Effects in Programmed Learning of a Verbal Concept. Journal of Educational Psychology. 1965, 56, 140-47.

Osborne, A. R. The Effects of Two Instructional Approaches on the Understanding of Subtraction by Grade Two Pupils. (Doctoral dissertation, University of Michigan) Ann Arbor, Michigan: University Microfilms, 1967, No. 67-8321.

Otis-Lennon Mental Ability Test, Elementary I, Form J. Harcourt, Brace and Jovanich, 1968. New York.

Overman, J. R. An Experimental Study of the Effect of the Method of Instruction on Transfer of Training in Arithmetic. Elementary School Journal. 1930, 31, 183-190.

Popham, W. J. and Baker, E. L. Rules for the Development of Instructional Products. In Baker, R. L. and Schutz, R. E. (eds.) Instructional Product Development. New York: Van Nostrand Reinhold, 1971.

Rector, R. E. and Henderson, K. B. The Relative Effectiveness of Four Strategies for Teaching Mathematical Concepts. Journal for Research in Mathematics Education, 1970, 1(2), 69-75.

Resnick, B. Issues in the Study of Learning Hierarchies. A paper presented at the 1970 annual meeting of the American Educational Research Association, Minneapolis, 1970.

Romberg, T. A., Fletcher, H. J. and Scott, J. A. A Measurement Approach to Elementary Mathematics Education. Working paper 12. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1968.

Romberg, T. A. and Harvey, J. G. Developing Mathematical Processes: Background and Projections. Working paper 14. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1969.

Romberg, T. A. and Planert, D. Developing Mathematical Processes: Pilot Tryout of Materials for Second Grade Children, Huegel

School 1968-69, Madison, Wisconsin. Working Paper No. 46. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1970.

Scheffé, H. The Analysis of Variance. New York: John Wiley and Sons, 1959.

Scott, J. A. The Effects on Short- and Long-Term Retention and on Transfer of Two Methods of Presenting Selected Geometry Concepts. Technical Report 138. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1970.

Short, J. and Haughey, B. Study of Sequencing Strategies, Proceedings of the 75th Annual Convention of the American Psychological Association. Washington: American Psychological Association, 1967, 317-318.

Silvern, L. S. Systems Engineering of Education I: The Evaluation of Systems Thinking in Education. Los Angeles: Educational and Training Consultants Company, 1965. (Library of Congress No. 65-27695).

Stanford Achievement Test, Primary I, Form X. Harcourt, Brace and World, 1964. New York.

Suppes, P., Jerman, M. and Brian, D. Computer-Assisted Instruction: Stanford's 1965-66 Program. New York: Academic Press, 1968.

Suppes, P. Stimulus Response Theory of Finite Automata. Journal of Mathematical Psychology, 1969, 8, 327-335.

Walbesser, H. H. An Evaluation Model and Its Application. Washington: American Association for the Advancement of Science, 1968.

Walbesser, H. H. and Eisenberg, T. A. A Review of Research on Behavioral Objectives and Learning Hierarchies. Mathematics Education Reports. Columbus, Ohio: ERIC Information Analysis Center for Science, Mathematics and Environmental Education, 1972.

Weaver, J. F. and Suydam, M. N. Meaningful Instruction in Mathematics Education. Mathematics Education Reports. Columbus, Ohio: ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, 1972.

Wiles, C. A., Romberg, T. A., and Moser, J. M. The Relative Effectiveness of Two Different Instructional Sequences Designed to Teach the Addition and Subtraction Algorithms. Technical Report 222. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1972.

Worthen, B. R. A Comparison of Discovery and Expository Sequencing in Elementary Mathematics Instruction. In Research in Mathematics Education. Washington: National Council of Teachers of Mathematics, 1967.

APPENDICES

Appendices B and C have been omitted from this publication, but are available on microfilm from Memorial Library, University of Wisconsin, Madison, Wisconsin.

APPENDIX A

Behaviors Associated with the Addition and Subtraction Algorithms for Two-Digit Whole Numbers

This Appendix contains the behaviors of the tasks A and S; they were derived from a larger hierarchy (see Harvey, Green, and McLeod, 1972). An identification of the power test variables with behaviors is also presented in this appendix.

Terminal Behavior AS - T (Addition-Subtraction)

Given a set of examples involving both addition and subtraction of 2 two-digit whole numbers, computes the required sums or differences using the appropriate algorithm. The sums are to be restricted to those less than 100.

Task A (Addition)

A1. Given the numeral phrase qb + sb, states the sentence qb + sb = (q + s)b.

A2. Given the numerals qb + r and sb + t in expanded notation, states the sentence

$$\begin{array}{llllll} q & b & + & r \\ s & b & + & t \\ \hline (q+s) & b & & +(r+t) \end{array}$$

A3. Given the numeral qb + r, where $r \geq b$, states the sentence qb + r = (q + 1)b + (r - b).

A4. Given the numerals qb + r and sb + t in expanded notation and having written the sentence (qb + r) + (sb + t) = (q + s)b + (r + t), finds numerals x and y such that (q + s)b + (r + t) = xb + y where y < b, and writes the sentence

$$\begin{array}{l} qb + r \\ sb + t \\ \hline xb + y \end{array}$$

A5. Given two two-digit numerals a and b written in compact notation, writes these numerals in expanded notation, finds their sum and the numeral c in compact notation which represents this sum, and writes the sentence a + b = c.

A6. Given two two-digit numerals a and b, written in compact notation, applies the Addition algorithm to a and b to find the numeral c in compact notation which represents this sum and writes the sentence a + b = c.

Task S (Subtraction)

S1. Given the numeral phrase qb - sb, where $q \geq s$, states the sentence qb - sb = (q - s)b.

S2. Given the numeral qb + r, states the sentence qb + r = (q - 1)b + (b + r).

S3. Given two numerals qb + r and sb + t, where $qb + r \geq sb + t$, compares r and t to determine whether $r \geq t$ or $r < t$.

S4. Given two numerals qb + r and sb + t, where qb + r > sb + t, and having determined that r < t, rewrites qb + r as (q - 1)b + (r + b), finds the differences (q - 1) - s and (r + b) - t, and writes the sentence

$$\begin{array}{rrcrr} q & b & + & r & \\ -\,s & b & - & t & \\ \hline (q-1-s) & b & + & (r+b-t) & \end{array}$$

S5. Given two numerals qb + r and sb + t, where $qb + r \geq sb + t$,

and having determined that $r \geq t$, finds the differences $q - s$ and $r - t$ and writes the sentence

$$\begin{array}{cccc} q & b & + & r \\ -\,s & b & - & t \\ \hline (q - s) & b & + & (r - t) \end{array}$$

S6. Given two two-digit numerals a and b written in compact notation, where $a \geq b$, writes these numerals in expanded notation, finds the difference $a - b$ and the numeral c in compact notation which represents this difference, and writes the sentence $a - b = c$.

S7. Given two two-digit numerals a and b written in compact notation, where $a \geq b$, applies the subtraction algorithm to the difference $a - b$ to find the numeral c in compact notation which represents this difference and writes the sentence $a - b = c$.

Entering Behaviors

E1. Given two two-digit numerals a and b written in compact notation, states that $a > b$ (or $b < a$).

E2. Given a numeral between 0 and 99 in compact notation, symbolically models this numeral in expanded notation.

E3. Given two numbers a and b whose sum is less than or equal to 20 (base 10), having represented the sum of a and b, and having found a numeral c in compact notation which is equal to the sum of a and b, writes the sentence $a + b = c$.

E4. Given two numbers a and b such that $0 \leq a < 20$, $0 \leq b < 10$, and $b \leq a$, having represented the difference a - b, and having found a numeral c in compact notation which is equal to that difference, writes the sentence a - b = c.

E5. Given two numbers a and b whose sum is less than or equal to 20 (base 10), states the compact number name for a + b and the complete sentence a + b = c.

E6. Given two numbers a and b such that $0 \leq a < 20$, $0 \leq b \leq 10$, and $b \leq a$, states the compact number name for a - b and the complete sentence a - b = c.

The Hierarchy of the Behaviors Associated With the Addition and Subtraction Algorithms.

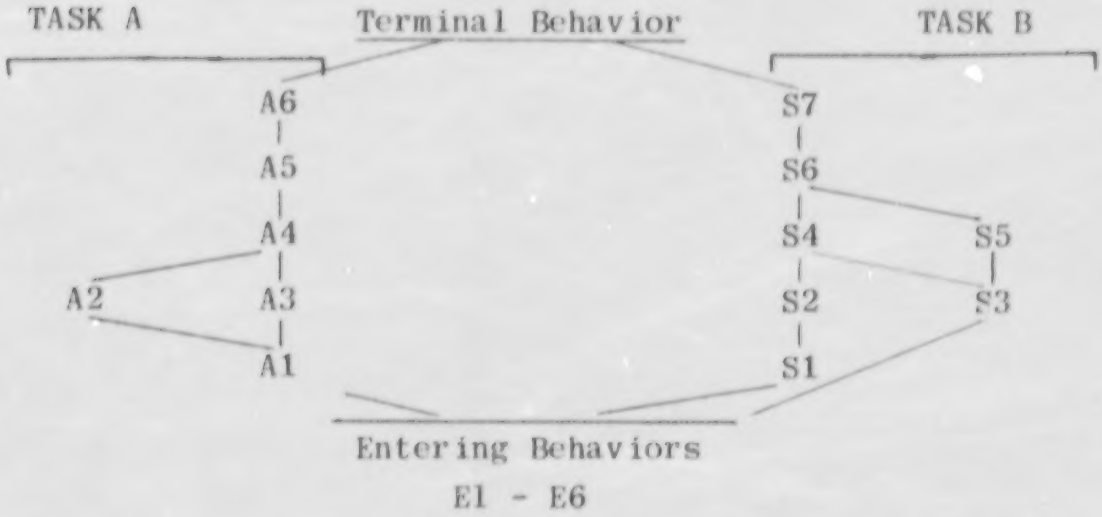


Table A.1

An Identification of Behaviors With Power Test Variables

Behaviors	Variables
A1	A(10)
A2	ANCE
A3	AR
A4	CE
A5	C + E, ANCE, CE
A6	ANCC, CC
S1	S(10)
S2	SR
S3	--
S4	BE
S5	SNBE
S6	C-E, SNBE, BE
S7	SNBC, BC
AS-T	ANCC, CC, SNBC, BC

APPENDIX D

Data on Subjects in the Experiment

This appendix contains the raw data that entered into the analyses reported in Chapter VIII.

Table D.1

Group AS Status Variables and OI Variables

	Status Variables					OI Variables					
Ss	IQ	Age	MET1	MET2	MET3	+F	-F	ANC	C	SNB	B
1	114	89	27	29	26	9	12	1	0	0	0
2	88	95	14	19	16	7	4	5	0	0	0
3	98	89	17	19	20	4	4	1	0	0	0
4	97	99	20	24	21	2	2	2	0	0	0
5	93	96	14	20	18	2	5	1	0	0	0
6	91	92	14	25	24	5	4	2	0	2	0
7	82	112	5	21	15	2	1	0	0	0	0
8*											
9	65	92	21	31	19	10	9	5	0	2	0
10*											
11	131	96	29	37	32	15	15	5	0	3	0
12	97	96	21	26	19	10	7	2	0	0	0
13	96	99	14	26	15	4	3	1	0	0	1
14	88	93	13	17	18	9	6	0	0	0	0
15	86	87	13	16	15	5	5	2	0	4	0
16*											
17	86	94	--	--	--	4	3	0	0	0	0
18	75	120	12	17	15	5	4	2	0	0	0

*Dropped from study

Table D.2

Group SA Status Variables and OI Variables

	Status Variables					OI Variables					
Ss	IQ	Age	MET1	MET2	MET3	+F	-F	ANC	C	SNB	B
1	121	92	25	34	30	10	5	3	0	1	0
2	104	97	22	32	23	4	2	2	0	0	0
3	98	97	21	27	27	7	8	1	3	0	0
4	97	99	14	20	17	4	6	4	0	1	0
5	94	95	21	29	21	7	9	4	0	1	0
6	90	89	--	--	--	4	4	0	0	0	0
7	82	115	17	25	18	6	8	4	0	5	0
8	79	100	9	20	11	2	1	0	0	0	0
9	68	104	16	10	16	3	4	2	0	1	0
10	64	100	11	19	15	1	2	0	0	0	0
11	112	89	15	28	18	6	6	3	0	0	0
12	98	96	17	23	23	4	2	0	0	0	0
13	94	93	18	19	17	3	8	4	0	3	0
14	93	93	15	22	20	4	3	0	0	0	0
15	87	98	17	24	17	6	3	0	0	0	0
16	83	105	10	23	19	5	4	1	0	2	0
17	80	98	13	18	13	3	3	3	0	1	0
18	66	106	8	16	10	0	1	0	0	0	0

Table D.3

Group I Status Variables and OI Variables

	Status Variables					OI Variables					
Ss	IQ	Age	MET1	MET2	MET3	+F	-F	ANC	C	SNB	B
1*											
2	102	94	10	21	18	2	2	0	0	0	0
3	101	93	10	25	17	3	1	1	0	0	0
4	96	94	13	27	16	5	6	1	0	2	0
5	96	92	--	--	--	1	1	1	0	2	0
6	91	89	16	19	17	3	2	1	2	0	0
7	85	97	12	23	14	5	4	0	0	0	0
8	75	95	22	26	19	7	7	2	0	2	0
9	75	110	12	12	15	3	3	1	0	0	0
10#											
11	105	91	22	15	17	10	8	5	0	5	0
12	97	96	25	33	26	10	7	4	1	5	0
13	93	92	20	24	20	3	4	2	0	2	0
14	93	90	16	20	10	3	2	0	0	0	0
15	89	99	13	19	16	5	5	0	0	0	0
16	82	99	23	24	23	9	7	2	0	1	0
17	81	98	8	12	9	2	1	1	0	0	0
18	77	92	11	26	18	5	3	2	0	1	0

*Dropped from study
#Not assigned

Table D.4

Group AS Observations P1 Through P6

	P1				P2				P3			
Sa	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1	1	0	0	0	2	0	2	0	--	--	--	--
2	2	0	0	0	2	0	0	0	2	0	0	0
3	1	0	0	0	--	--	--	--	--	--	--	--
4	2	0	2	0	2	0	0	0	1	0	2	0
5	0	1	0	0	1	0	0	0	1	1	0	0
6	2	0	1	0	2	0	1	0	2	0	2	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8*												
9	2	1	2	0	2	1	1	0	2	0	2	0
10*												
11	2	1	2	0	2	0	2	0	2	0	2	0
12	--	--	--	--	--	--	--	--	2	0	0	0
13	--	--	--	--	--	--	--	--	2	0	0	0
14	0	1	0	0	2	0	1	0	1	0	2	0
15	1	0	2	0	2	0	2	0	1	0	2	0
16*												
17	0	0	0	0	--	--	--	--	0	0	0	0
18	0	0	1	0	2	0	2	0	1	1	0	0

	P4				P5				P6			
1	2	1	2	0	2	0	2	0	2	2	0	0
2	2	2	0	0	2	1	2	0	2	2	2	0
3	2	0	1	0	2	1	2	0	2	2	2	0
4	2	2	2	0	2	2	1	0	2	1	1	0
5	2	0	0	0	1	1	0	0	1	2	0	0
6	2	0	2	0	2	2	2	0	2	1	1	0
7	0	0	0	0	--	--	--	--	0	0	0	0
8*												
9	2	2	2	0	2	2	2	0	2	2	2	0
10*												
11	2	2	2	0	2	2	2	0	2	2	2	0
12	1	2	1	0	2	0	2	0	0	2	1	0
13	2	1	0	0	2	0	0	0	1	2	0	0
14	2	0	2	0	2	0	2	0	2	1	0	0
15	--	--	--	--	--	--	--	--	--	--	--	--
16*												
17	0	0	0	0	0	0	0	0	0	0	0	0
18	2	1	2	0	2	1	0	0	2	1	2	0

* Dropped from study

Table D.5

Group AS Observations P7 Through P12

	P7				P8				P9			
Ss	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1	1	1	2	0	2	2	2	0	2	2	1	0
2	--	--	--	--	--	--	--	--	--	--	--	--
3	2	2	0	0	--	--	--	--	--	--	--	--
4	2	2	2	0	2	2	2	0	2	1	2	0
5	2	2	0	0	2	2	1	0	2	0	0	0
6	2	2	2	0	2	2	1	0	2	2	2	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8*												
9	2	2	2	0	2	1	2	1	2	2	2	0
10*												
11	2	2	2	0	2	2	2	0	2	2	2	0
12	2	0	2	0	2	0	2	0	2	0	2	0
13	--	--	--	--	2	2	2	0	1	0	2	0
14	2	0	0	0	0	1	1	0	1	0	1	0
15	2	0	2	0	1	1	1	0	2	1	1	0
16*												
17	--	--	--	--	--	--	--	--	0	0	0	0
18	1	0	1	0	2	1	1	0	2	0	1	0

	P10				P11				P12			
1	1	2	1	1	2	1	2	2	2	2	2	2
2	2	1	2	0	2	1	2	0	2	1	0	0
3	--	--	--	--	--	--	--	--	--	--	--	--
4	2	1	2	1	2	2	2	2	2	2	2	2
5	1	0	0	0	2	2	1	0	2	2	0	0
6	2	2	2	0	2	1	2	0	2	2	2	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8*												
9	2	2	2	2	2	2	2	2	2	2	2	2
10*	1	0	0	0	2	2	1	0	2	2	0	0
11	2	2	2	2	2	2	2	2	2	2	2	2
12	2	0	2	0	2	0	2	1	2	0	2	2
13	2	1	0	0	2	2	1	0	0	1	1	0
14	2	1	2	1	2	1	1	1	0	0	2	0
15	2	1	2	0	2	0	2	0	2	1	2	0
16*												
17	0	0	0	0	0	0	0	0	0	0	0	0
18	2	2	2	0	2	1	0	0	2	1	2	0

* Dropped from study

Table D.6

Group SA Observations P1 Through P6

	P1				P2				P3			
Ss	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1	--	--	--	--	2	0	1	0	2	1	2	0
2	2	0	1	0	2	0	1	0	2	0	2	0
3	1	0	0	0	1	2	0	0	2	2	0	0
4	2	0	0	0	1	1	1	0	1	0	1	0
5	1	0	0	0	2	0	1	0	2	1	2	0
6	--	--	--	--	--	--	--	--	0	0	0	0
7	1	0	1	0	2	0	2	0	2	0	2	0
8	0	0	0	0	1	0	0	0	0	0	0	0
9	2	0	0	0	2	0	2	0	2	0	1	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	1	1	0	1	0	1	0	1	0	1	0
12	0	0	1	0	2	0	2	0	2	0	2	0
13	1	0	2	0	2	0	2	0	2	0	2	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	--	--	--	--	0	0	0	0	0	0	0	0
16	2	0	1	0	1	0	1	0	1	0	1	0
17	1	0	1	0	--	--	--	--	1	0	0	0
18	0	0	0	0	2	0	1	0	2	0	0	0

	P4				P5				P6			
1	1	0	0	0	1	1	2	1	1	2	2	0
2	2	0	2	0	1	1	1	0	2	0	2	0
3	1	1	0	0	1	1	0	0	0	0	2	0
4	2	0	2	0	2	0	2	0	1	0	2	0
5	2	0	2	0	2	1	2	0	2	1	2	0
6	--	--	--	--	0	0	0	0	2	0	1	0
7	2	0	2	0	2	0	1	0	2	1	1	0
8	0	0	0	0	--	--	--	--	--	--	--	--
9	2	0	0	0	2	0	2	0	2	0	2	0
10	0	0	0	0	--	--	--	--	0	0	0	0
11	2	0	2	0	1	0	2	1	2	0	2	2
12	2	0	2	0	2	0	2	0	2	0	0	0
13	1	0	2	0	2	0	2	0	2	0	2	0
14	0	0	0	0	0	0	0	0	0	0	1	0
15	--	--	--	--	0	0	0	0	0	0	0	0
16	2	0	0	0	1	0	1	0	1	0	2	0
17	2	0	1	0	--	--	--	--	--	--	--	--
18	1	0	0	0	1	0	1	0	2	0	0	0

Table D.7

Groups SA Observations P7 Through P12

	P7				P8				P9			
Ss	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1	2	0	2	1	2	1	2	1	2	2	1	1
2	2	0	2	0	1	2	2	0	2	2	2	1
3	2	0	2	0	1	2	1	0	--	--	--	--
4	2	0	1	1	2	0	1	0	2	1	2	0
5	1	2	2	1	1	1	1	0	2	1	2	1
6	0	0	0	0	0	0	0	0	0	1	0	0
7	0	0	1	2	2	2	2	2	2	2	2	2
8	0	0	0	0	0	0	0	0	0	0	0	0
9	2	0	1	0	2	0	1	0	2	0	2	0
10	1	0	0	0	0	0	0	0	0	0	0	0
11	2	0	2	2	2	2	2	1	2	2	2	2
12	2	1	1	0	2	0	2	0	1	1	1	0
13	2	0	2	1	2	1	2	0	2	1	1	1
14	0	0	0	0	--	--	--	--	0	0	0	0
15	0	0	0	0	2	0	0	0	2	0	2	0
16	1	0	0	0	1	0	1	0	0	0	0	0
17	--	--	--	--	--	--	--	--	--	--	--	--
18	0	0	0	0	1	0	1	0	0	0	0	0
	P10				P11				P12			
1	2	2	2	2	2	1	1	1	2	2	2	2
2	2	2	2	0	2	2	2	1	2	2	2	2
3	--	--	--	--	1	0	1	0	1	1	1	0
4	2	1	1	0	1	0	1	0	1	0	2	0
5	2	1	2	0	2	2	1	0	2	2	2	0
6	1	0	0	0	1	1	0	0	2	1	0	0
7	2	1	1	2	2	2	2	2	2	1	2	2
8	1	1	0	0	0	0	0	0	0	0	0	0
9	2	1	2	0	2	0	1	0	2	2	1	0
10	0	1	0	0	2	0	0	0	2	0	1	0
11	2	2	2	2	2	2	2	0	2	2	2	1
12	1	1	1	1	2	2	1	2	2	2	2	0
13	2	2	2	1	2	2	2	1	2	2	2	1
14	0	0	0	0	0	0	0	0	0	0	0	0
15	2	2	2	0	2	2	2	0	--	--	--	--
16	1	1	0	0	2	0	1	0	1	1	1	0
17	--	--	--	--	--	--	--	--	--	--	--	--
18	1	0	2	0	--	--	--	--	1	0	1	0

Table D.8

Group I Observations P1 Through P6

	P1				P2				P3			
Ss	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1*												
2	0	0	0	0	0	0	0	0	1	1	0	0
3	1	0	0	0	1	0	2	0	2	0	2	0
4	1	0	0	0	0	0	0	0	1	1	2	0
5	1	0	2	0	2	0	1	0	--	--	--	--
6	0	0	0	0	--	--	--	--	1	0	0	0
7	0	0	0	0	0	1	0	0	0	0	0	0
8	2	0	2	0	2	0	2	0	2	0	2	0
9	1	0	1	0	1	0	0	0	0	0	0	0
10#												
11	2	0	2	0	2	0	2	0	2	0	2	0
12	1	2	2	0	2	1	2	0	2	2	2	0
13	2	0	2	0	2	0	2	0	2	1	1	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0
16	2	0	1	0	--	--	--	--	1	0	0	0
17	1	0	0	0	0	0	0	0	2	0	0	0
18	2	1	2	0	0	0	2	0	2	0	1	0

	P4				P5				P6			
1*												
2	0	0	0	0	2	0	0	0	2	0	1	0
3	1	0	0	0	2	0	0	0	1	0	0	0
4	2	0	2	0	2	0	2	0	2	0	1	0
5	--	--	--	--	1	1	2	0	2	0	2	0
6	1	1	0	0	1	1	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
8	2	0	2	0	--	--	--	--	2	0	1	0
9	0	0	1	0	2	0	1	0	1	0	2	0
10#												
11	2	0	2	0	2	0	2	0	2	0	2	0
12	2	1	1	0	2	2	2	1	2	2	2	0
13	2	0	2	0	2	1	2	1	2	1	2	1
14	--	--	--	--	1	0	1	0	2	0	0	0
15	--	--	--	--	0	1	1	0	0	0	0	0
16	2	0	1	0	2	0	1	0	--	--	--	--
17	1	0	0	0	1	0	0	0	1	0	0	0
18	2	0	0	0	--	--	--	--	--	--	--	--

* Dropped from study
\# Not assigned

Table D.9

Group I Observations P7 Through P12

	P7				P8				P9			
Ss	ANC	C	SNB	B	ANC	C	SNB	B	ANC	C	SNB	B
1*												
2	2	0	2	0	--	--	--	--	--	--	--	--
3	2	0	0	0	2	0	2	0	2	0	0	0
4	2	0	2	0	2	2	2	0	2	2	2	1
5	1	0	2	0	2	1	2	0	2	2	2	1
6	0	0	0	0	2	1	2	0	0	1	2	0
7	0	0	0	0	2	1	0	0	0	1	0	0
8	2	0	1	0	2	2	2	0	2	2	2	1
9	1	0	2	0	1	0	2	0	2	0	2	0
10#												
11	2	0	2	0	2	1	2	1	2	1	2	1
12	2	2	2	0	2	1	2	2	2	2	2	1
13	2	0	2	0	2	2	2	0	1	2	1	1
14	1	0	1	0	--	--	--	--	--	--	--	--
15	0	1	0	1	2	1	1	0	1	1	1	0
16	1	0	1	0	2	2	2	1	2	2	2	1
17	1	0	0	0	2	0	0	0	0	0	0	0
18	1	2	1	0	2	0	1	0	2	1	2	0

	P10				P11				P12			
1*												
2	--	--	--	--	--	--	--	--	--	--	--	--
3	2	0	2	0	1	0	0	0	2	0	2	1
4	2	1	2	1	2	1	1	1	2	2	2	2
5	2	2	2	1	2	1	2	0	2	2	2	2
6	2	1	2	1	1	1	2	1	2	1	2	0
7	1	0	1	0	2	0	1	0	2	0	1	0
8	2	2	2	2	--	--	--	--	2	2	2	2
9	2	0	1	0	2	0	2	0	2	0	1	0
10#												
11	2	2	2	2	2	2	2	2	2	1	2	2
12	2	2	2	2	0	2	2	2	2	2	2	2
13	2	2	2	2	2	2	2	0	2	2	2	2
14	--	--	--	--	2	0	0	0	2	0	0	0
15	2	2	1	0	1	2	1	0	2	2	0	0
16	2	2	2	2	2	0	2	1	2	2	2	2
17	1	0	0	0	1	0	0	0	0	0	0	0
18	2	1	2	0	0	2	1	1	2	1	2	0

* Dropped from study
\# Not assigned

Table D.10

Group AS Observation 02

Ss	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1	9	9	4	8	4	6	3	3	3	3	3	4
2	8	7	4	8	4	0	0	0	4	2	4	0
3	--	--	--	--	--	--	--	--	--	--	--	--
4	12	12	4	6	4	6	0	3	4	4	4	2
5	--	--	--	--	--	--	--	--	--	--	--	--
6	8	5	2	3	2	4	1	0	4	1	4	3
7	4	3	0	0	0	0	0	0	0	0	0	0
8*												
9	10	9	5	10	5	10	0	3	4	4	4	4
10*	--	--	--	--	--	--	--	--	--	--	--	--
11	19	20	5	10	5	9	4	4	4	4	4	3
12	7	7	3	0	2	3	0	1	4	0	4	4
13	2	1	0	0	1	2	0	0	0	1	3	1
14	10	7	1	0	1	4	0	0	4	0	4	2
15	5	2	3	0	4	0	0	0	4	3	4	0
16*												
17	5	0	0	0	0	0	0	0	0	0	0	0
18	4	2	1	3	1	0	0	0	0	0	1	0

Ss	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1	4	4	4	4	4	4	4	3
2	4	4	4	4	3	0	4	4
3	--	--	--	--	--	--	--	--
4	4	4	4	3	4	3	4	4
5	--	--	--	--	--	--	--	--
6	4	4	4	4	3	3	4	3
7	2	4	0	0	0	0	0	0
8*								
9	4	4	4	4	4	4	4	4
10*								
11	4	4	4	4	4	4	4	4
12	4	4	3	0	4	4	4	4
13	4	4	3	1	4	3	4	4
14	4	4	4	0	3	3	4	4
15	4	4	4	0	4	0	4	4
16*								
17	4	4	0	0	0	0	0	0
18	3	4	4	3	3	0	3	3

* Dropped from study

Table D.11

Group SA Observation 02

Sa	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1	10	9	2	3	2	2	4	3	4	3	3	4
2	5	7	4	7	5	4	2	3	4	4	4	2
3	9	9	1	3	1	1	0	0	3	0	1	0
4	4	2	3	6	3	0	0	0	3	4	4	0
5	9	10	4	7	5	0	0	0	4	4	2	0
6	2	3	2	1	2	0	0	0	4	3	2	0
7	6	5	3	5	3	6	0	1	4	4	4	3
8	3	1	0	1	0	0	0	0	0	0	0	0
9	6	1	2	0	1	0	0	0	3	3	1	0
10	3	1	3	5	0	0	0	0	4	0	0	0
11	4	5	4	7	4	2	1	3	4	4	3	2
12	4	3	3	2	2	1	0	0	3	3	2	2
13	4	3	2	4	1	0	0	0	4	4	4	4
14	--	--	--	--	--	--	--	--	--	--	--	--
15	3	4	2	3	2	0	0	0	4	3	4	0
16	4	0	2	4	2	0	1	0	1	0	2	0
17	--	--	--	--	--	--	--	--	--	--	--	--
18	3	1	2	0	1	0	1	0	0	0	0	0

Sa	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1	4	4	4	3	4	4	4	4
2	4	4	4	4	4	4	4	4
3	4	4	4	0	3	0	4	4
4	4	4	4	4	4	0	4	4
5	4	4	4	4	4	1	4	3
6	4	4	4	0	4	0	4	4
7	4	4	4	2	4	4	4	4
8	2	2	1	0	1	0	1	0
9	4	4	4	1	4	0	4	4
10	2	2	4	3	1	0	4	1
11	4	4	4	4	4	3	4	4
12	4	4	4	4	4	2	4	3
13	4	4	4	4	2	0	4	3
14	--	--	--	--	--	--	--	--
15	4	4	4	3	4	0	4	4
16	4	4	3	3	4	1	4	4
17	--	--	--	--	--	--	--	--
18	4	4	3	2	1	0	3	2

Table D.12

Group I Observation 02

Sa	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1*												
2	--	--	--	--	--	--	--	--	--	--	--	--
3	5	1	0	2	2	1	0	0	1	0	1	1
4	4	3	2	4	2	2	0	0	1	1	3	0
5	5	5	3	5	3	0	3	4	4	3	2	2
6	4	2	2	4	2	2	0	0	4	4	4	2
7	2	3	1	0	0	0	0	0	0	0	0	0
8	9	7	4	8	5	7	0	0	4	3	3	3
9	1	1	2	0	2	0	0	0	3	0	0	0
10#												
11	10	9	4	8	5	8	0	0	3	4	3	3
12	13	6	4	6	4	4	1	0	4	4	3	3
13	4	3	2	4	1	4	0	0	4	4	4	4
14	3	4	1	0	1	0	0	0	4	0	0	0
15	4	2	3	4	3	0	1	0	3	3	2	0
16	6	7	4	6	4	7	1	0	4	4	4	4
17	1	0	1	0	0	0	0	0	1	0	0	0
18	2	0	1	3	0	3	0	0	3	2	1	2

Sa	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1*								
2	--	--	--	--	--	--	--	--
3	0	0	2	1	0	0	0	0
4	4	4	4	4	4	4	3	4
5	4	4	4	4	4	3	0	0
6	4	4	4	3	4	2	4	4
7	4	4	0	0	2	0	3	3
8	4	4	4	4	4	3	4	4
9	1	0	4	0	3	0	0	0
10#								
11	4	4	4	4	4	4	4	4
12	4	4	4	4	4	4	4	4
13	4	3	3	4	3	3	4	4
14	4	4	4	0	1	0	4	2
15	4	4	4	4	3	0	4	3
16	4	4	4	4	4	4	4	4
17	4	4	0	0	0	0	3	0
18	4	4	3	2	3	3	4	4

* Dropped from study
\# Not assigned

Table D.13

Group AS Observation 03

Ss	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1	12	14	4	10	5	8	4	4	3	3	4	4
2	10	8	5	9	3	0	0	0	4	4	4	0
3	--	--	--	--	--	--	--	--	--	--	--	--
4	10	10	3	4	4	2	0	0	3	3	4	2
5	11	4	1	1	1	0	0	0	3	3	0	0
6	--	--	--	--	--	--	--	--	--	--	--	--
7	2	3	0	0	0	0	0	0	0	0	0	0
8*												
9	10	7	5	9	5	10	0	0	4	4	4	4
10*												
11	23	22	5	10	5	10	4	4	4	4	4	4
12	8	8	4	0	3	2	0	0	4	0	4	4
13	5	5	2	0	2	1	0	0	3	4	3	4
14	5	5	2	4	2	3	0	1	4	3	3	2
15	5	3	2	5	3	0	0	0	4	4	4	0
16*												
17	5	0	0	0	0	0	0	0	2	2	0	0
18	6	4	2	2	2	1	0	0	2	2	3	0

Ss	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1	4	4	4	3	4	2	4	3
2	4	2	4	4	3	0	4	4
3	--	--	--	--	--	--	--	--
4	4	4	4	3	4	3	4	4
5	4	2	2	0	0	0	2	0
6	--	--	--	--	--	--	--	--
7	4	4	0	0	0	0	0	0
8*								
9	4	4	4	4	4	2	4	4
10*								
11	4	3	4	4	4	3	3	4
12	4	4	4	0	4	4	4	4
13	4	4	4	4	3	4	4	4
14	4	4	4	3	3	2	4	4
15	4	4	4	0	4	0	4	4
16*								
17	4	4	3	0	0	0	2	0
18	4	4	4	4	3	0	4	4

* Dropped from study

Table D.14

Group SA Observation 03

Ss	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1	11	10	2	6	2	4	4	4	4	4	4	4
2	7	9	4	6	4	5	2	3	4	2	4	4
3	10	10	2	1	1	0	0	0	3	0	1	0
4	5	3	4	7	1	0	1	0	3	4	0	0
5	--	--	--	--	--	--	--	--	--	--	--	--
6	5	2	2	0	2	0	0	0	4	1	4	0
7	5	6	3	6	3	5	1	0	4	4	4	4
8	3	0	0	0	0	0	0	0	1	0	0	0
9	6	1	4	2	0	0	0	0	0	0	0	0
10	5	1	2	5	0	0	0	0	4	0	0	0
11	9	7	2	3	2	4	4	2	4	4	4	4
12	7	5	1	2	1	3	0	0	3	2	1	1
13	6	2	5	5	1	0	0	0	3	3	3	1
14	6	8	3	0	2	0	0	0	3	3	4	0
15	4	4	2	3	1	0	0	0	4	4	4	0
16	4	1	3	3	1	0	0	0	1	0	1	0
17	--	--	--	--	--	--	--	--	--	--	--	--
18	3	4	0	0	1	0	0	0	0	0	0	0

Ss	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1	4	4	3	4	4	4	4	3
2	4	4	4	3	4	2	4	3
3	4	4	2	1	4	0	4	4
4	4	4	4	3	3	0	4	4
5	--	--	--	--	--	--	--	--
6	4	4	4	0	4	0	4	4
7	4	4	4	4	4	4	4	4
8	2	2	0	0	0	0	0	0
9	4	4	4	3	3	0	4	4
10	2	4	4	3	0	0	3	0
11	4	4	4	4	4	4	4	4
12	4	3	3	3	2	2	3	4
13	4	4	4	4	4	0	4	3
14	4	4	4	0	2	0	3	4
15	4	4	4	3	4	0	4	4
16	3	4	3	3	0	0	3	3
17	--	--	--	--	--	--	--	--
18	4	4	3	0	2	0	3	2

Table D.15

Group I Observation 03

Ss	+F	-F	ANC	C	SNB	B	AR	SR	ANCE	CE	SNBE	BE
1*												
2	--	--	--	--	--	--	--	--	--	--	--	--
3	4	2	1	0	1	0	0	0	1	2	1	0
4	6	7	2	4	2	4	0	0	4	3	4	1
5	5	4	5	6	3	0	0	0	4	3	3	0
6	6	6	2	5	1	3	0	0	4	4	4	2
7	2	3	1	0	2	0	0	0	1	0	1	0
8	6	6	4	7	5	6	1	2	4	3	4	3
9	2	5	1	0	1	0	0	1	4	0	3	0
10#												
11	10	9	4	7	4	6	0	0	4	4	3	4
12	9	8	4	5	5	7	0	0	3	4	4	1
13	5	6	2	5	2	4	0	0	4	4	4	3
14	1	5	1	3	1	0	0	0	4	1	0	0
15	--	--	--	--	--	--	--	--	--	--	--	--
16	2	3	2	1	2	0	0	0	3	4	4	3
17	1	0	1	0	1	0	0	0	2	0	0	0
18	2	2	2	2	2	1	0	0	4	2	3	1

Ss	C+E	C-E	ANCC	CC	SNBC	BC	A(10)	S(10)
1*								
2	--	--	--	--	--	--	--	--
3	2	4	1	0	0	0	0	1
4	4	4	4	3	3	4	4	4
5	4	4	4		4	1	0	0
6	4	4	4	4	4	0	4	3
7	4	4	4	0	4	0	4	4
8	4	4	4	4	4	4	4	4
9	4	4	3	0	4	0	3	4
10#								
11	4	4	4	3	4	3	4	4
12	4	4	4	3	4	4	4	4
13	4	4	4	4	4	3	3	4
14	4	4	2	0	0	0	4	2
15	--	--	--	--	--	--	--	--
16	3	4	4	4	4	4	4	4
17	3	4	2	0	2	0	4	0
18	4	4	4	3	2	3	2	4

*Dropped from study
#Not assigned

U.S.Government Printing Office: 1974 — 750-487/1036 Region 5-1

National Evaluation Committee

Helen Bain
Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
University of Colorado

Sue Buel
Dissemination and Installation Services
Northwest Regional Educational Laboratory

Francis S. Chase
Professor Emeritus
University of Chicago

George E. Dickson
College of Education
University of Toledo

Chester W. Harris
Graduate School of Education
University of California

Hugh J. Scott
Consultant
National Evaluation Committee

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Joanna Williams
Professor of Psychology and Education
Columbia University

Executive Committee

William R. Bush
Director, Program Planning and Management
Deputy Director, R & D Center

M. Vere DeVault
Professor
School of Education

Herbert J. Klausmeier
Principal Investigator
R & D Center

Joel R. Levin
Principal Investigator
R & D Center

Donald N. McIsaac
Associate Dean, School of Education
University of Wisconsin

Richard A. Rossmiller, Committee Chairman
Director
R & D Center

Len VanEss
Associate Vice Chancellor
University of Wisconsin–Madison

Dan Woolpert
Director, Management Systems
R & D Center

Faculty of Principal Investigators

Vernon L. Allen
Professor
Psychology

B. Dean Bowles
Associate Professor
Educational Administration

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
V. A. C. Henmon Professor
Educational Psychology

Gisela Labouvie
Assistant Professor
Educational Psychology

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

James Lipham
Professor
Educational Administration

Wayne Otto
Professor
Curriculum and Instruction

Robert Petzold
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Dennis W. Spuck
Assistant Professor
Educational Administration

Richard L. Venezky
Associate Professor
Computer Science

Larry M. Wilder
Assistant Professor
Communication Arts